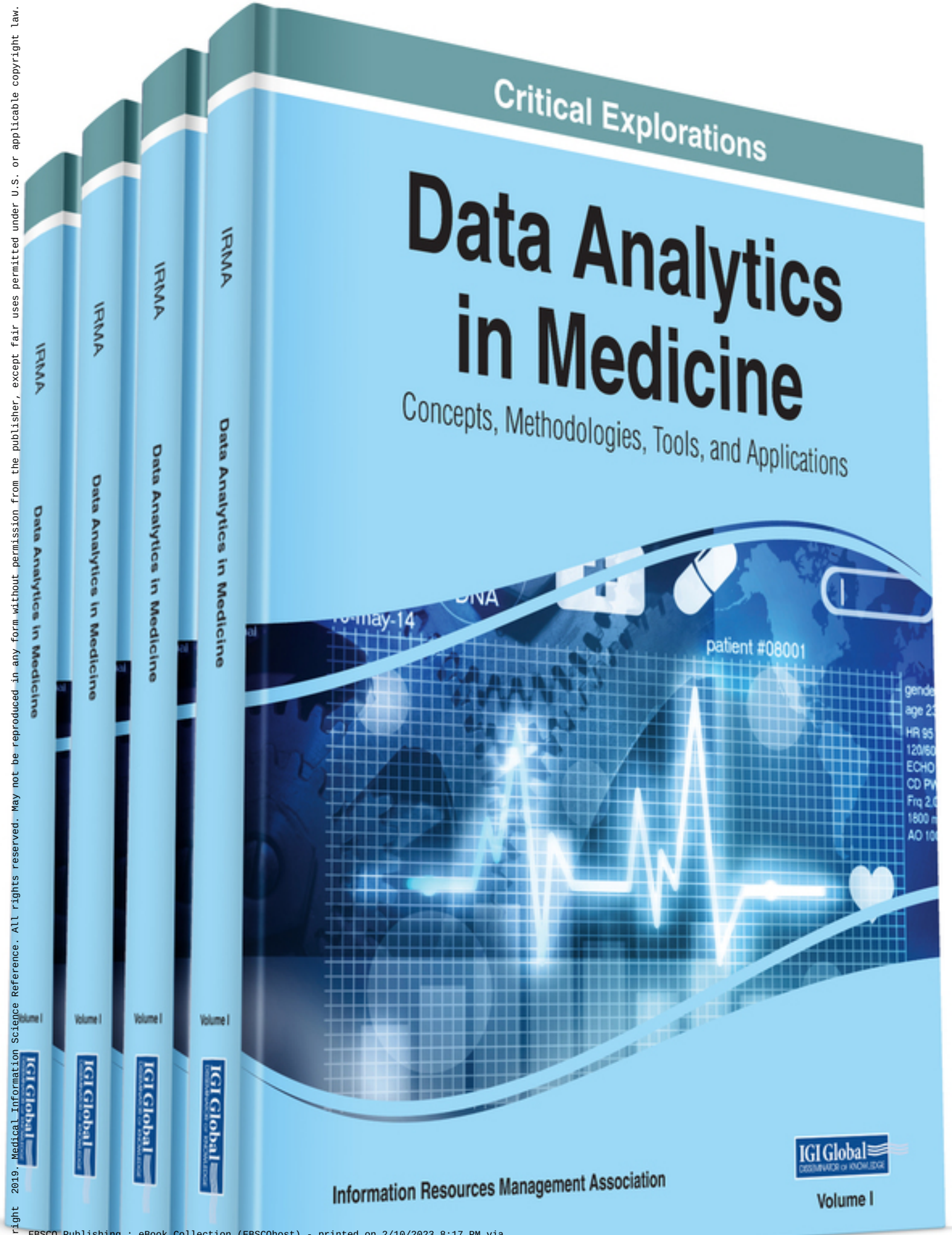


Copyright 2019, Medical Information Science Reference. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 2/10/2023 8:17 PM via
AN: 2318624 ; Information Resources Management Association.; Data Analytics in Medicine : Concepts, Methodologies, Tools, and Applications
Account: ns335141

Data Analytics in Medicine:

Concepts, Methodologies, Tools, and Applications

Information Resources Management Association
USA

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Published in the United States of America by
IGI Global
Medical Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Copyright © 2020 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Names: Information Resources Management Association, editor.
Title: Data analytics in medicine : concepts, methodologies, tools, and applications / Information Resources Management Association, editor.
Description: Hershey, PA : Medical Information Science Reference, [2020] | Includes bibliographical references and index. | Summary: “”This book examines practical applications of healthcare analytics for improved patient care, resource allocation, and medical performance, as well as for diagnosing, predicting, and identifying at-risk populations”-- Provided by publisher.
Identifiers: LCCN 2019027003 (print) | LCCN 2019027004 (ebook) | ISBN 9781799812043 (h/c) | ISBN 9781799812050 (eISBN)
Subjects: MESH: Medical Informatics Applications | Data Science--methods | Big Data | Data Mining
Classification: LCC R855.3 (print) | LCC R855.3 (ebook) | NLM W 26.5 | DDC 610.285--dc23
LC record available at https://lccn.loc.gov/2019027003
LC ebook record available at https://lccn.loc.gov/2019027004

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: eresources@igi-global.com.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Editor-in-Chief

Mehdi Khosrow-Pour, DBA
Information Resources Management Association, USA

Associate Editors

Steve Clarke, *University of Hull, UK*
Murray E. Jennex, *San Diego State University, USA*
Ari-Veikko Anttiroiko, *University of Tampere, Finland*

Editorial Advisory Board

Sherif Kamel, *American University in Cairo, Egypt*
In Lee, *Western Illinois University, USA*
Jerzy Kisielnicki, *Warsaw University, Poland*
Amar Gupta, *Arizona University, USA*
Craig van Slyke, *University of Central Florida, USA*
John Wang, *Montclair State University, USA*
Vishanth Weerakkody, *Brunel University, UK*

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

List of Contributors

Abautret, Frank / *Collège Max Jacob, France* 898
Abelha, António / *University of Minho, Portugal* 865
Abuelyaman, Eltayeb Salih / *University of Dammam, Saudi Arabia* 1989
Ahamad, Shaik Shakeel / *Majmaah University, Kingdom of Saudi Arabia* 437
Akaichi, Jalel / *King Khalid University, Saudi Arabia* 268
Alamri, Abdullah / *University of Jeddah, Saudi Arabia* 169
Alberga, Domenico / *Università degli Studi di Bari "Aldo Moro", Italy* 194
Al-Hakim, Latif / *University of South Australia, Australia* 1416
Ali, A. / *Ain Shams University, Egypt* 1052
Almakadmeh, Khaled / *Hashemite University, Jordan* 1341
Alsaadi, Mahmood / *Princess Sumaya University for Technology, Jordan* 1341
Alsaghier, Hisham M. / *Majmaah University, Kingdom of Saudi Arabia* 437
Al-Ssulami, Abdulrakeeb M. / *King Saud University, Saudi Arabia* 298
Amaral, Francisco Emanuel Batista / *University of Azores, Portugal* 1193
Anandan, Kavitha / *SSN College of Engineering, India* 1654
Anandaram, Harishchander / *Sathyabama University, India* 1875
Antoniou, Panagiotis / *Aristotle University of Thessaloniki, Greece* 1963
Apolinario-Arzube, Oscar / *University of Guayaquil, Ecuador* 1023
Arafah, Mohammed Amer / *King Saud University, Saudi Arabia* 298
Archondakis, Stavros / *401 Military Hospital, Greece* 1921
Arshad, Arfan / *International Islamic University, Malaysia & Jazan University, Kingdom of Saudi Arabia* 150
Bagchi, Angshuman / *University of Kalyani, India* 1895
Baha, Nadia / *University of Science and Technology USTHB, Algeria* 726
Bahar, Nurhidayah / *UCSI University, Malaysia* 340
Bahri, Shamshul / *University of Malaya, Malaysia* 340
Balathayil, Divya / *SSN College of Engineering, India* 1654
Ballet, Pascal / *Université de Bretagne Occidentale, France* 898
Bamidis, Panagiotis / *Aristotle University of Thessaloniki, Greece* 1963
Banerjee, Anwesha / *Jadavpur University, India* 833
Barik, Rabindra K. / *KIIT University, India* 1826
Basharat, Zarrin / *Fatima Jinnah Women University, Pakistan* 1556
Bennett, Kevin J. / *University of South Carolina, USA* 1038
Berler, Alexander / *IHE – Europe, Belgium* 357
Bharadwaj, Charan / *Sri Jayachamarajendra College of Engineering, India* 410
Bhartiya, Shalini / *IITM, India* 1485

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bhattacharya, Sylvia / *Georgia Southern University, USA* 2002
Bhattacharyya, Pushpak / *Indian Institute of Technology Patna, India* 1502
Bibi, Monazza / *Fatima Jinnah Women University, Pakistan* 1556
Billis, Antonis S. / *Aristotle University of Thessaloniki, Greece* 1963
Bisoi, R. / *Siksha 'O' Anusandhan University, India* 779
Borkloe, Foster / *Catholic University College of Ghana, Ghana* 117
Boudriga, Noureddine / *University of Carthage, Tunisia* 1781
Bourquard, Karima / *IHE – Europe, France* 357
Briggs, Jim / *University of Portsmouth, UK* 879
Brown, Patrick / *Thermo Scientific NanoDrop Products, USA* 1364
Brox, Ellen / *NORUT, Norway* 1599
Cabitza, Federico / *Università degli Studi di Milano-Bicocca, Italy* 1148
Chakravarty, S. / *Orissa Engineering College, India* 779
Chavan, Ankita Arun / *M. S. Ramaiah Institute of Technology, India* 288
Chen, Edward T. / *University of Massachusetts Lowell, USA* 1858, 1943
Chengaiyan, Sandhya / *SSN College of Engineering, India* 1654
Chowdhury, Linkon / *Chittagong University of Engineering and Technology, Bangladesh* 1168
Cook, Jack S. / *The College at Brockport, USA* 1758
Costa, Paulo / *George Mason University, USA* 1222
Costarides, Vassilia / *National Technical University of Athens, Greece* 2033
Cvetković, Mario / *University of Split, Croatia* 249
Daily, Sandra / *Duke Medical Center, USA* 1738
Darko-Lartey, Richard / *Catholic University College of Ghana, Ghana* 117
Das, Amit / *University of Kalyani, India* 1895
Dasgupta, Rakhi / *University of Kalyani, India* 1895
Dash, Nilamadhab / *C. V. Raman College of Engineering, India* 502
Dash, P. K. / *Siksha 'O' Anusandhan University, India* 779
Dash, Sujata / *North Orissa University, India* 645, 707
Demurjian, Steven A. / *University of Connecticut, USA* 528, 1254
Dev, Kaushik / *Chittagong University of Engineering and Technology, Bangladesh* 1168
Dey, Nilanjan / *Techno India College of Technology, India* 1168
Đogaš, Zoran / *University of Split, Croatia* 249
Dong, MingChui / *University of Macau, Macau* 663
Dorić, Vicko / *University of Split, Croatia* 249
Doster, Barbara / *University of Alabama, USA* 1738
Drissi, Khalil El Khamlichi / *Blaise Pascal University, France* 249
Duarte, Júlio / *University of Minho, Portugal* 865
Dubey, Harishchandra / *The University of Texas at Dallas, USA* 1826
Durst, Carolin / *University of Bremen, Germany* 1
Ekbal, Asif / *Indian Institute of Technology Patna, India* 1502
El Alaoui, Said Ouatik / *Sidi Mohammed Ben Abdellah University, Morocco* 103
El Guemhioui, Karim / *University of Quebec in Outaouais, Canada* 528
Ellingsen, Gunnar / *Arctic University of Norway, Norway* 1704
Ellouze, Nourhene / *University of Carthage, Tunisia* 1781
Elsayad, Dina / *Ain Shams University, Egypt* 1052
Evertsen, Gunn / *NORUT, Norway* 1599

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fathalla, Said / *Bonn University, Germany & Alexandria University, Egypt* 471, 1814
Fitch, Tineke / *University of Portsmouth, UK* 879
Frantzidıs, Christos A. / *Aristotle University of Thessaloniki, Greece* 1963
Fronville, Alexandra / *Université de Bretagne Occidentale, France* 898
Fu, Meng / *University of Melbourne, Australia* 852
Gaivéo, José / *Polytechnic Institute of Setubal, Portugal* 1683
Galli, Brian J. / *Long Island University – Post, USA* 1467
Garza, Roberto / *Escuela de Ingenieria y Tecnologias de Informacion, Mexico* 614
Gewald, Corinna / *Neu-Ulm University of Applied Sciences, Germany* 1720
Gewald, Heiko / *Neu-Ulm University of Applied Sciences, Germany* 1720
Ghosh, Poulami / *Jadavpur University, India* 833
Giokas, Kostas / *National Technical University of Athens, Greece* 2033
Girdhar, Anup / *Sedulity Groups, India* 1485
Girish, Amoolya / *M. S. Ramaiah Institute of Technology, India* 1328
Gomes, Jorge / *Universidade de Lisboa, Portugal* 2018
Gómez, Lorena G / *Escuela de Ingenieria y Tecnologias de Informacion, Mexico* 614
Gonzalez, Rafael A. / *Pontificia Universidad Javeriana, Colombia* 325
Good, Alice / *University of Portsmouth, UK* 879
Gürsel, Güney / *Command post of Gendarme Logistics, School ol of Technical and Auxiliary Forces, Turkey* 50
Gyaase, Patrick Ohemeng / *Catholic University College of Ghana, Ghana* 117
Hahn, Michael E. / *University of Oregon, USA* 762
Hallberg, Josef / *Luleå University of Technology, Sweden* 1599
Hassanien, Aboul Ella / *Cairo University, Egypt* 565
Haueisen, Jens / *Technical University Ilmenau, Germany* 249
Helm, Emmanuel / *University of Applied Sciences Upper Austria, Austria* 424
Hertzberg, Libi / *Tel Aviv University, Israel & Weizmann Institute of Science, Israel* 580
Hirche, Johannes / *Luleå University of Technology, Sweden* 1599
Hung, Che-Lun / *Providence University, Taiwan* 629
Idrish, Sherina / *North South University, Bangladesh* 1632
Iqbal, Mehree / *North South University, Bangladesh* 1632
Issa, Mohamed / *Zagazig University, Egypt* 565
Jácome, Cristina / *University of Aveiro, Portugal* 815
Jeong, Hyeon-Cheol / *Sahmyook University, Korea* 1674
Jeyanathan, Josephine / *Kalasalingam University, India* 1328
Jiang, Furen / *Changchun University, China* 1006
Joseph, Rhoda / *Penn State University, USA* 1364
Joshi, Deepak / *Indian Institute of Technology (IIT) Delhi, India* 762
Kamal, Md Sarwar / *East West University, Bangladesh* 1168
Kamalashree, N. / *Sri Jayachamarajendra College of Engineering, India* 410
Kannot, Yaman M. Khalid / *Arab Academy for Science, Egypt* 1814
Karakitsos, Petros / *National and Kapodistrian University of Athens, Greece* 1921
Kartsidis, Panagiotis Emmanouil / *Aristotle University of Thessaloniki, Greece* 1963
Katehakis, Dimitrios G. / *Institute of Computer Science (ICS), Greece* 1450
Kelder, Jo-Anne / *University of Tasmania, Tasmania* 1537
Kesari, Anupama / *Sri Jayachamarajendra College of Engineering, India* 410

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Khan, Mohammad Ibrahim / *Chittagong University of Engineering and Technology, Bangladesh* 1168
Konstantinidis, Evdokimos / *Aristotle University of Thessaloniki, Greece* 1963
Konstantinidis, Stathis Th. / *NORUT, Norway* 1599
Koumaditis, Konstantinos / *University of Piraeus, Greece, & Aarhus University, Denmark* 133
Koutsouris, Dimitris / *National Technical University of Athens, Greece* 2033
Krauss, Oliver / *University of Applied Sciences Upper Austria, Austria* 424
Kueng, Josef / *Johannes Kepler University Linz, Austria* 424
Kumar, Dilip / *SLIET, India* 743
Kumar, Rajeev / *DAVIET, India* 743
Kummervold, Per Egil / *NORUT, Norway* 1599
Labbadi, Wissem / *University of Tunis, Tunisia* 268
Labbé, Cyril / *University Grenoble Alpes, France* 325
Lace-Costigan, Gemma / *University of Salford, UK* 1283
Lagos-Ortiz, Katty / *University of Guayaquil, Ecuador* 1023
Lakaraju, Sandeep / *Dakota State University, USA* 1378
Lara-Diaz, Victor J. / *Escuela de Ingenieria y Tecnologias de Informacion, Mexico* 614
Lavariega, Juan C. / *Escuela de Ingenieria y Tecnologias de Informacion, Mexico* 614
Lazakidou, Athina / *University of Peloponnese, Greece* 1522
Leonetti, Francesco / *Università degli Studi di Bari "Aldo Moro", Italy* 194
Lewis, Carmen / *Troy University, USA* 1738
Lin, Chun-Yuan / *Chang Gung University, Taiwan* 629
Locoro, Angela / *Università degli Studi di Milano-Bicocca, Italy* 1148
López, Elvis O. / *Centro Brasileiro de Pesquisas Físicas, Brazil* 1076
Luna-Aveiga, Harry / *University of Guayaquil, Ecuador* 1023
Ma, JiaLi / *University of Macau, Macau* 663
Machado, José / *University of Minho, Portugal* 865
Mallia, Anthony J. / *Edmond Scientific Company, USA* 2048
Mangiatordi, Giuseppe F. / *Università degli Studi di Bari "Aldo Moro", Italy* 194
Mann, Joshua / *University of Mississippi Medical Center, USA* 1038
Mareels, Iven / *University of Melbourne, Australia* 852
Margari, Niki / *National and Kapodistrian University of Athens, Greece* 1921
Marques, Alda / *University of Aveiro, Portugal* 815
Mathkour, Hassan / *King Saud University, Saudi Arabia* 298
Mazumder, Ankita / *Jadavpur University, India* 833
Medina-Moreira, José / *University of Guayaquil, Ecuador* 1023
Mehrotra, Deepti / *Amity University, India* 1485
Menaka, M. / *Indira Gandhi Centre for Atomic Research, India* 1328
Merrouche, Fairouz / *University of Science and Technology USTHB, Algeria* 726
Mishra, Brojo Kishore / *C. V. Raman College of Engineering, India* 502, 1826
Misra, Rachita / *C. V. Raman College of Engineering, India* 502
Montagna, Sara / *Università di Bologna, Italy* 369
Moumtzoglou, Anastasius / *Hellenic Society for Quality and Safety in Healthcare, Greece, & P. & A. Kyriakou Children's Hospital, Greece* 74
Mukhtar, Wafaa Faisal / *Sudan University of Science and Technology, Sudan* 1989
Mukred, Muaadh / *National University of Malaysia (UKM), Malaysia* 1578

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Múnera, Alejandro Sierra / *Pontificia Universidad Javeriana, Colombia* 325
Murali, Leema / *M. S. Ramaiah Institute of Technology, India* 1328
Neely, M. Pamela / *The College at Brockport, USA* 1758
Ngqila, Kholekile Hazel / *Walter Sisulu University, South Africa* 986
Nicolotti, Orazio / *University of Bari, Italy* 194
Nisha, Nabila / *North South University, Bangladesh* 1632
Noor, Talal / *Taibah University, Saudi Arabia* 182
Noordin, Mohamad Fauzan / *International Islamic University, Malaysia* 150
O'Donnell, Eileen / *Technological University Dublin, Ireland* 13
O'Donnell, Liam / *Technological University Dublin, Ireland* 13
Okuboyejo, Senanu R. / *Covenant University, Nigeria* 487
Oliveira, Ana / *University of Aveiro, Portugal* 815
Oliveira, Daniela / *University of Minho, Portugal* 865
Omicini, Andrea / *Università di Bologna, Italy* 369
Omisade, Omobolanle / *University of Portsmouth, UK* 879
Or, Calvin / *The University of Hong Kong, China* 1297
Othman, Roslina / *International Islamic University, Malaysia* 150
Ouyang, Lijing / *Centers for Disease Control & Prevention, USA* 1038
Pal, Randip / *Texas Tech University, USA* 226
Pandey, Anukul / *Dr. B. R. Ambedkar National Institute of Technology, India* 930
Panigrahi, Chhabi / *University of Rajastan, India* 1826
Paraskevopoulos, Evangelos / *Aristotle University of Thessaloniki, Greece* 1963
Pasmanik-Chor, Metsada / *Tel Aviv University, Israel* 580
Passeri, Giovanna Ilaria / *Molecular Discovery Ltd., UK* 194
Pathak, Parth S / *ezDI, LLC, India* 1502
Patra, Bichitrananda / *KMBB College of Engineering and Technology, India* 645
Pedersen, Rune / *University Hospital North Norway, Norway* 1704
Peña, Angel Alberto Garcia / *Pontificia Universidad Javeriana, Colombia* 325
Pichavant, Karine / *Université de Bretagne Occidentale, France* 898
Pinho, Cátia / *University of Aveiro, Portugal* 815
Poljak, Dragan / *University of Split, Croatia* 249
Pomares-Quimbaya, Alexandra / *Pontificia Universidad Javeriana, Colombia* 325
Pothet, Alain / *Académie de Créteil, France* 898
Pouliakis, Abraham / *National and Kapodistrian University of Athens, Greece* 74, 1921
Priyadarshini, Rojalina / *C. V. Raman College of Engineering, India* 502, 1826
Pullen, Troy / *Queensland State Archives, Australia* 1416
Qasaimeh, Malik / *Princess Sumaya University for Technology, Jordan* 1341
Qureshi, Basit / *Prince Sultan University, Saudi Arabia* 593
Raj, Nitika / *Sri Jayachamarajendra College of Engineering, India* 410
Rajasekaran, Gowri / *Periyar University, India* 1306
Rath, Mamata / *C.V. Raman College of Engineering, India* 1401, 1839
Rathipriya R / *Periyar University, India* 1306
Ravi, Prabha / *M. S. Ramaiah Institute of Technology, India* 1328
Rawat, Danda B. / *Georgia Southern University, USA* 2002
Rekhis, Slim / *University of Carthage, Tunisia* 1781
Riechert, Jana / *University of Erlangen-Nuremberg, Germany* 1

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rifat, Afrin / *North South University, Bangladesh* 1632
Rivière, Jérémy / *Université de Bretagne Occidentale, France* 898
Rodin, Vincent / *Université de Bretagne Occidentale, France* 898
Rodrigues, João Manuel / *University of Aveiro, Portugal* 815
Rodríguez, Julián Camilo Daza / *Pontificia Universidad Javeriana, Colombia* 325
Romão, Mário / *Universidade de Lisboa, Portugal* 2018
Ryan, Jim / *Auburn University at Montgomery, USA* 1738
Saha, Sriparna / *Indian Institute of Technology Patna, India* 1502
Saini, Barjinder Singh / *Dr. B. R. Ambedkar National Institute of Technology, India* 930
Salas-Zárate, María del Pilar / *University of Murcia, Spain* 1023
Sánchez, Yaira K. Rivera / *University of Connecticut, USA* 1254
Sarrouti, Mourad / *Sidi Mohammed Ben Abdellah University, Morocco* 103
Schuler, Andreas / *University of Applied Sciences Upper Austria, Austria* 424
Sehgal, Manika / *Institute of Microbial Technology, India* 551
Sekar, Booma Devi / *University of Macau, Macau* 663
Shalini, Priya / *Sri Jayachamarajendra College of Engineering, India* 410
Shankar, Prabhu RV / *University of California Davis, USA* 410
Shaqrah, Amin / *Taibah University, Saudi Arabia* 182
Shedeed, Howida A. / *Ain Shams University, Egypt* 1052
Shenbagavalli, A. / *National Engineering College, India* 1328
Shivakumar, Manu / *State University of New York at Buffalo, USA* 410
Silva, Sandra / *Unidade de Saúde de Ilha de S. Miguel Azores, Portugal* 1193
Silva-Cavazos, Manuel J. / *Escuela de Ingenieria y Tecnologias de Informacion, Mexico* 614
Singh, Butta / *Guru Nanak Dev University Regional Campus, India* 930
Singh, TirathaRaj / *Jaypee University of Information Technology, India* 551
Singla, Tony / *DAVIET, India* 743
Siragusa, Lydia / *Molecular Discovery Ltd., UK* 194
Sirur, Manjunath / *M. S. Ramaiah Institute of Technology, India* 1328
Skafidas, Stan / *University of Melbourne, Australia* 852
Sood, Neetu / *Dr. B. R. Ambedkar National Institute of Technology, India* 930
Srinivas, Sowrabha / *Sri Jayachamarajendra College of Engineering, India* 410
Srinivas, Sushmitha / *M. S. Ramaiah Institute of Technology, India* 1328
Sriraam, Natarajan / *M. S. Ramaiah Institute of Technology, India* 288, 1328
Styliadis, Charis / *Aristotle University of Thessaloniki, Greece* 1963
Su, Min / *Beijing University of Technology, China* 629
Swatman, Paula M / *Swinburne University, Australia* 1537
Taborda, Jaime Andrés Pérez / *Universidad Tecnológica de Pereira, Colombia & Instituto de Microelectrónica de Madrid-CSIC, Colombia* 1076
Tagadur, Nagabhushana N. / *Sri Jayachamarajendra College of Engineering, India* 410
Tan, Jianjun / *Beijing University of Technology, China* 629
Tedmori, Sara / *Princess Sumaya University for Technology, Jordan* 1341
Tejaswini, S. / *M. S. Ramaiah Institute of Technology, India* 288
Themistocleous, Marinos / *University of Piraeus, Greece & University of Coimbra, Portugal* 133
Theron, Michaël / *Université de Bretagne Occidentale, France* 898
Thomas, Christy Bobby / *M.S.Ramaiah University of Applied Sciences, India* 1654
Tiago, Flávio / *UAC, Portugal* 1193

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tiago, Maria Teresa Borges / *University of Azores, Portugal* 1193
Tibarewala, D.N. / *Jadavpur University, India* 833
Tiwari, Basant / *Hawassa University, Ethiopia* 456
Tiwari, Vivek / *IIIT-NR, India* 456
Tolba, Mohamed F. / *Ain Shams University, Egypt* 1052
Traxler, Barbara / *University of Applied Sciences Upper Austria, Austria* 424
Trisciuzzi, Daniela / *Università degli Studi di Bari "Aldo Moro", Italy* 194
Tsaloukidis, Nikolaos / *University of Peloponnese, Greece* 1522
Tsoromokos, Dimitrios / *University of Peloponnese, Greece* 1522
Ulle, Anand Raj / *Sri Jayachamarajendra College of Engineering, India* 410
Ulriksen, Gro-Hilde / *University Hospital North Norway, Norway* 1704
Unnithan, Chandana / *Torrens University Australia, Australia* 1537
Valencia-García, Rafael / *University of Murcia, Spain* 1023
Vasanthamani, Saranya / *Sri Krishna College of Engineering and Technology, India* 971
Velandia, Oscar Mauricio Muñoz / *Pontificia Universidad Javeriana, Colombia* 325
Venkataraman, B. / *Indira Gandhi Centre for Atomic Research, India* 1328
Vidaković, Maja Rogić / *University of Split, Croatia* 249
Walczak, Steven / *University of South Florida, USA* 390, 487
Wang, Chung-Hung / *Chang Gung University, Taiwan* 629
Wang, Yong / *Dakota State University, USA* 1378
Wickramasinghe, Nilmini / *Epworth HealthCare, Australia & Deakin University, Australia* 1
Wijesekera, Duminda / *George Mason University, USA* 1222
William, Harrison / *Catholic University College of Ghana, Ghana* 117
Xefteris, Stefanos / *Aristotle University of Thessaloniki, Greece* 1963
Xiao, Zhiguo / *Changchun University, China* 1006
Xu, Dianxiang / *Boise State University, USA* 1378
Yadav, Shweta / *Indian Institute of Technology Patna, India* 1502
Yang, Yongji / *Changchun University, China* 1006
Yasmin, Azra / *Fatima Jinnah Women University, Pakistan* 1556
Ye, Jin / *Chang Gung University, Taiwan* 629
Yitzhaky, Assif / *Weizmann Institute of Science, Israel* 580
Yousefi, Mohammadmahdi Rezaei / *The Ohio State University, USA* 1110
Yu, Bo / *George Mason University, USA* 1222
Yusof, Zawiyah M. / *National University of Malaysia (UKM), Malaysia* 1578
Zarakovitis, Dimitrios / *University of Peloponnese, Greece* 1522
Zulim, Ivana / *University of Split, Croatia* 249
Zygomalas, Apollon / *University of Patras, Greece* 2033

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table of Contents

Preface xxvii

Volume I

Section 1
Fundamental Concepts and Theories

Chapter 1
A Guideline to Use Activity Theory for Collaborative Healthcare Information Systems Design 1
Carolin Durst, University of Bremen, Germany
Nilmini Wickramasinghe, Epworth HealthCare, Australia & Deakin University, Australia
Jana Riechert, University of Erlangen-Nuremberg, Germany

Chapter 2
A Review of Type 1 Diabetes (T1D): To Assist Patients' and Carers' Understanding of This Condition 13
Eileen O'Donnell, Technological University Dublin, Ireland
Liam O'Donnell, Technological University Dublin, Ireland

Chapter 3
For Better Healthcare Mining Health Data 50
Güney Gürsel, Command post of Gendarme Logistics, School ol of Technical and Auxiliary Forces, Turkey

Chapter 4
Population Health Management and the Science of Individuality 74
Anastasius Moumtzoglou, Hellenic Society for Quality and Safety in Healthcare, Greece, & P. & A. Kyriakou Children's Hospital, Greece
Abraham Pouliakis, National and Kapodistrian University of Athens, Greece

Section 2
Development and Design Methodologies

Chapter 5
A Yes/No Answer Generator Based on Sentiment-Word Scores in Biomedical Question Answering 103
Mourad Sarrouti, Sidi Mohammed Ben Abdellah University, Morocco
Said Ouatik El Alaoui, Sidi Mohammed Ben Abdellah University, Morocco

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 6
Towards an Integrated Electronic Medical Records System for Quality Healthcare in Ghana: An Exploratory Factor Analysis 117
Patrick Ohemeng Gyaase, Catholic University College of Ghana, Ghana
Richard Darko-Lartey, Catholic University College of Ghana, Ghana
Harrison William, Catholic University College of Ghana, Ghana
Foster Borkloe, Catholic University College of Ghana, Ghana

Chapter 7
SOA Governance in Healthcare: Beyond Early Ideas to a Structured Framework 133
Konstantinos Koumaditis, University of Piraeus, Greece, & Aarhus University, Denmark
Marinos Themistocleous, University of Piraeus, Greece & University of Coimbra, Portugal

Chapter 8
KIIPF - An Integrated Inter-Operable Knowledge Management Process Framework for Healthcare: Implementation in Pakistani Healthcare Industry 150
Mohamad Fauzan Noordin, International Islamic University, Malaysia
Arfan Arshad, International Islamic University, Malaysia & Jazan University, Kingdom of Saudi Arabia
Roslina Othman, International Islamic University, Malaysia

Chapter 9
Semantic Health Mediation and Access Control Manager for Interoperability Among Healthcare Systems 169
Abdullah Alamri, University of Jeddah, Saudi Arabia

Chapter 10
A Conceptual Framework for an Extension Access Control Models in Saudi Arabia Healthcare Systems 182
Amin Shaqrah, Taibah University, Saudi Arabia
Talal Noor, Taibah University, Saudi Arabia

Chapter 11
Strategies of Virtual Screening in Medicinal Chemistry 194
Giovanna Ilaria Passeri, Molecular Discovery Ltd., UK
Daniela Trisciuzzi, Università degli Studi di Bari "Aldo Moro", Italy
Domenico Alberga, Università degli Studi di Bari "Aldo Moro", Italy
Lydia Siragusa, Molecular Discovery Ltd., UK
Francesco Leonetti, Università degli Studi di Bari "Aldo Moro", Italy
Giuseppe F. Mangiatordi, Università degli Studi di Bari "Aldo Moro", Italy
Orazio Nicolotti, University of Bari, Italy

Chapter 12
Relationships between Models of Genetic Regulatory Networks with Emphasis on Discrete State Stochastic Models 226
Randip Pal, Texas Tech University, USA

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 13
Integral Equation Formulations and Related Numerical Solution Methods in Some Biomedical Applications of Electromagnetic Fields: Transcranial Magnetic Stimulation (TMS), Nerve Fiber Stimulation 249
Dragan Poljak, University of Split, Croatia
Mario Cvetković, University of Split, Croatia
Vicko Dorić, University of Split, Croatia
Ivana Zulim, University of Split, Croatia
Zoran Đogaš, University of Split, Croatia
Maja Rogić Vidaković, University of Split, Croatia
Jens Haueisen, Technical University Ilmenau, Germany
Khalil El Khamlichi Drissi, Blaise Pascal University, France

Chapter 14
Efficient Algorithm for Answering Fuzzy Medical Requests in Pervasive Healthcare Information Systems 268
Wissem Labbadi, University of Tunis, Tunisia
Jalel Akaichi, King Khalid University, Saudi Arabia

Chapter 15
Development of Portable Medical Electronic Device for Infant Cry Recognition: A Primitive Experimental Study 288
Natarajan Sriraam, M. S. Ramaiah Institute of Technology, India
S. Tejaswini, M. S. Ramaiah Institute of Technology, India
Ankita Arun Chavan, M. S. Ramaiah Institute of Technology, India

Chapter 16
Efficient String Matching Algorithm for Searching Large DNA and Binary Texts 298
Abdulrakeeb M. Al-Ssulami, King Saud University, Saudi Arabia
Hassan Mathkour, King Saud University, Saudi Arabia
Mohammed Amer Arafah, King Saud University, Saudi Arabia

Chapter 17
Concept Attribute Labeling and Context-Aware Named Entity Recognition in Electronic Health Records 325
Alexandra Pomares-Quimbaya, Pontificia Universidad Javeriana, Colombia
Rafael A. Gonzalez, Pontificia Universidad Javeriana, Colombia
Oscar Mauricio Muñoz Velandia, Pontificia Universidad Javeriana, Colombia
Angel Alberto Garcia Peña, Pontificia Universidad Javeriana, Colombia
Julián Camilo Daza Rodríguez, Pontificia Universidad Javeriana, Colombia
Alejandro Sierra Múnera, Pontificia Universidad Javeriana, Colombia
Cyril Labbé, University Grenoble Alpes, France

Chapter 18
Knowledge Management Process-Oriented Strategy for Healthcare Organizations 340
Nurhidayah Bahar, UCSI University, Malaysia
Shamshul Bahri, University of Malaya, Malaysia

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 19
Use-Case Driven Approach for a Pragmatic Implementation of Interoperability in eHealth 357
Karima Bourquard, IHE – Europe, France
Alexander Berler, IHE – Europe, Belgium

Chapter 20
Agent-Based Modelling in Multicellular Systems Biology 369
Sara Montagna, Università di Bologna, Italy
Andrea Omicini, Università di Bologna, Italy

Chapter 21
The Role of Artificial Intelligence in Clinical Decision Support Systems and a Classification Framework 390
Steven Walczak, University of South Florida, USA

Chapter 22
Predictive Modeling of Surgical Site Infections Using Sparse Laboratory Data 410
Prabhu RV Shankar, University of California Davis, USA
Anupama Kesari, Sri Jayachamarajendra College of Engineering, India
Priya Shalini, Sri Jayachamarajendra College of Engineering, India
N. Kamalashree, Sri Jayachamarajendra College of Engineering, India
Charan Bharadwaj, Sri Jayachamarajendra College of Engineering, India
Nitika Raj, Sri Jayachamarajendra College of Engineering, India
Sowrabha Srinivas, Sri Jayachamarajendra College of Engineering, India
Manu Shivakumar, State University of New York at Buffalo, USA
Anand Raj Ulle, Sri Jayachamarajendra College of Engineering, India
Nagabhushana N. Tagadur, Sri Jayachamarajendra College of Engineering, India

Chapter 23
Towards Semantic Interoperability in Health Data Management Facilitating Process Mining 424
Barbara Traxler, University of Applied Sciences Upper Austria, Austria
Emmanuel Helm, University of Applied Sciences Upper Austria, Austria
Oliver Krauss, University of Applied Sciences Upper Austria, Austria
Andreas Schuler, University of Applied Sciences Upper Austria, Austria
Josef Kueng, Johannes Kepler University Linz, Austria

Chapter 24
A Secure Robust and Privacy Enhanced Mobile Healthcare Framework 437
Hisham M. Alsaghier, Majmaah University, Kingdom of Saudi Arabia
Shaik Shakeel Ahamad, Majmaah University, Kingdom of Saudi Arabia

Chapter 25
An Intelligent Multi-Objective Framework of Pervasive Information Computing 456
Basant Tiwari, Hawassa University, Ethiopia
Vivek Tiwari, IIIT-NR, India

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 3
Tools and Technologies

Chapter 26
Detecting Human Diseases Relatedness: A Spreading Activation Approach Over Ontologies 471
Said Fathalla, Bonn University, Germany & Alexandria University, Egypt

Chapter 27
An Artificial Neural Network Classification of Prescription Nonadherence 487
Steven Walczak, University of South Florida, USA
Senanu R. Okuboyejo, Covenant University, Nigeria

Volume II

Chapter 28
Applying CI in Biology through PSO 502
Rojalina Priyadarshini, C. V. Raman College of Engineering, India
Nilamadhab Dash, C. V. Raman College of Engineering, India
Brojo Kishore Mishra, C. V. Raman College of Engineering, India
Rachita Misra, C. V. Raman College of Engineering, India

Chapter 29
Semantic Reconciliation of Electronic Health Records Using Semantic Web Technologies 528
Karim El Guemhioui, University of Quebec in Outaouais, Canada
Steven A. Demurjian, University of Connecticut, USA

Chapter 30
Principles and Analysis of Biological Networks: Biological Pathways and Network Motifs 551
Manika Sehgal, Institute of Microbial Technology, India
TirathaRaj Singh, Jaypee University of Information Technology, India

Chapter 31
Multiple Sequence Alignment Optimization Using Meta-Heuristic Techniques 565
Mohamed Issa, Zagazig University, Egypt
Aboul Ella Hassanien, Cairo University, Egypt

Chapter 32
GEView (Gene Expression View) Tool for Intuitive and High Accessible Visualization of Expression Data for Non-Programmer Biologists 580
Libi Hertzberg, Tel Aviv University, Israel & Weizmann Institute of Science, Israel
Assif Yitzhaky, Weizmann Institute of Science, Israel
Metsada Pasmanik-Chor, Tel Aviv University, Israel

Chapter 33
An Affordable Hybrid Cloud Based Cluster for Secure Health Informatics Research 593
Basit Qureshi, Prince Sultan University, Saudi Arabia

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 34
EEMI - An Electronic Health Record for Pediatricians: Adoption Barriers, Services and Use in Mexico 614
Juan C. Lavariega, Escuela de Ingenieria y Tecnologias de Informacion, Mexico
Roberto Garza, Escuela de Ingenieria y Tecnologias de Informacion, Mexico
Lorena G Gómez, Escuela de Ingenieria y Tecnologias de Informacion, Mexico
Victor J. Lara-Diaz, Escuela de Ingenieria y Tecnologias de Informacion, Mexico
Manuel J. Silva-Cavazos, Escuela de Ingenieria y Tecnologias de Informacion, Mexico

Chapter 35
Constructing a Bioinformatics Platform with Web and Mobile Services Based on NVIDIA Jetson TK1 629
Chun-Yuan Lin, Chang Gung University, Taiwan
Jin Ye, Chang Gung University, Taiwan
Che-Lun Hung, Providence University, Taiwan
Chung-Hung Wang, Chang Gung University, Taiwan
Min Su, Beijing University of Technology, China
Jianjun Tan, Beijing University of Technology, China

Chapter 36
Genetic Diagnosis of Cancer by Evolutionary Fuzzy-Rough based Neural-Network Ensemble 645
Sujata Dash, North Orissa University, India
Bichitrananda Patra, KMBB College of Engineering and Technology, India

Chapter 37
Wired and Wireless Distributed e-Home Healthcare System 663
Booma Devi Sekar, University of Macau, Macau
JiaLi Ma, University of Macau, Macau
MingChui Dong, University of Macau, Macau

Chapter 38
Hybrid Ensemble Learning Methods for Classification of Microarray Data: RotBagg Ensemble Based Classification 707
Sujata Dash, North Orissa University, India

Chapter 39
Fall Detection Depth-Based Using Tilt Angle and Shape Deformation 726
Fairouz Merrouche, University of Science and Technology USTHB, Algeria
Nadia Baha, University of Science and Technology USTHB, Algeria

Chapter 40
Energy Efficient Particle Optimized Compressed ECG Data over Zigbee Environment 743
Dilip Kumar, SLIET, India
Rajeev Kumar, DAVIET, India
Tony Singla, DAVIET, India

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 41
Electromyogram and Inertial Sensor Signal Processing in Locomotion and Transition Classification 762
Deepak Joshi, Indian Institute of Technology (IIT) Delhi, India
Michael E. Hahn, University of Oregon, USA

Chapter 42
A Hybrid Kernel Extreme Learning Machine and Improved Cat Swarm Optimization for Microarray Medical Data Classification 779
S. Chakravarty, Orissa Engineering College, India
R. Bisoi, Siksha 'O' Anusandhan University, India
P. K. Dash, Siksha 'O' Anusandhan University, India

Chapter 43
Integrated Approach for Automatic Crackle Detection Based on Fractal Dimension and Box Filtering 815
Cátia Pinho, University of Aveiro, Portugal
Ana Oliveira, University of Aveiro, Portugal
Cristina Jácome, University of Aveiro, Portugal
João Manuel Rodrigues, University of Aveiro, Portugal
Alda Marques, University of Aveiro, Portugal

Chapter 44
Electromyography-Based Functional Electrical Stimulation (FES) in Rehabilitation 833
Poulami Ghosh, Jadavpur University, India
Ankita Mazumder, Jadavpur University, India
Anwesha Banerjee, Jadavpur University, India
D.N. Tibarewala, Jadavpur University, India

Chapter 45
A Novel Delay-Based GFSK Demodulator in 65 nm CMOS for Low Power Biomedical Applications 852
Meng Fu, University of Melbourne, Australia
Stan Skafidas, University of Melbourne, Australia
Iven Mareels, University of Melbourne, Australia

Chapter 46
Step Towards Interoperability in Nursing Practice 865
Daniela Oliveira, University of Minho, Portugal
Júlio Duarte, University of Minho, Portugal
António Abelha, University of Minho, Portugal
José Machado, University of Minho, Portugal

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 47
An Analysis of Factors Affecting Postnatal Depression Intervention Adherence 879
Omobolanle Omisade, University of Portsmouth, UK
Alice Good, University of Portsmouth, UK
Tineke Fitch, University of Portsmouth, UK
Jim Briggs, University of Portsmouth, UK

Chapter 48
Modelling and Simulating Complex Systems in Biology: Introducing NetBioDyn – A Pedagogical and Intuitive Agent-Based Software..... 898
Pascal Ballet, Université de Bretagne Occidentale, France
Jérémy Rivière, Université de Bretagne Occidentale, France
Alain Pothet, Académie de Créteil, France
Michaël Theron, Université de Bretagne Occidentale, France
Karine Pichavant, Université de Bretagne Occidentale, France
Frank Abautret, Collège Max Jacob, France
Alexandra Fronville, Université de Bretagne Occidentale, France
Vincent Rodin, Université de Bretagne Occidentale, France

Chapter 49
Analysis of Electrocardiogram Data Compression Techniques: A MATLAB-Based Approach 930
Anukul Pandey, Dr. B. R. Ambedkar National Institute of Technology, India
Barjinder Singh Saini, Dr. B. R. Ambedkar National Institute of Technology, India
Butta Singh, Guru Nanak Dev University Regional Campus, India
Neetu Sood, Dr. B. R. Ambedkar National Institute of Technology, India

Chapter 50
A Study on Lifetime Enhancement and Reliability in Wearable Wireless Body Area Networks 971
Saranya Vasanthamani, Sri Krishna College of Engineering and Technology, India

Section 4
Utilization and Applications

Chapter 51
Pluralistic Tendencies in Healing Abantu Illness in the Contemporary South Africa 986
Kholekile Hazel Ngqila, Walter Sisulu University, South Africa

Chapter 52
Kinect-Based Limb Rehabilitation Methods 1006
Yongji Yang, Changchun University, China
Zhiguo Xiao, Changchun University, China
Furen Jiang, Changchun University, China

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Volume III

Chapter 53
Knowledge Acquisition Through Ontologies from Medical Natural Language Texts 1023
José Medina-Moreira, University of Guayaquil, Ecuador
Katty Lagos-Ortiz, University of Guayaquil, Ecuador
Harry Luna-Aveiga, University of Guayaquil, Ecuador
Oscar Apolinario-Arzube, University of Guayaquil, Ecuador
María del Pilar Salas-Zárate, University of Murcia, Spain
Rafael Valencia-García, University of Murcia, Spain

Chapter 54
Utilizing Combined Claims and Clinical Datasets for Research Among Potential Cases of Rare Diseases..... 1038
Kevin J. Bennett, University of South Carolina, USA
Joshua Mann, University of Mississippi Medical Center, USA
Lijing Ouyang, Centers for Disease Control & Prevention, USA

Chapter 55
PAGeneRN: Parallel Architecture for Gene Regulatory Network 1052
Dina Elsayad, Ain Shams University, Egypt
A. Ali, Ain Shams University, Egypt
Howida A. Shedeed, Ain Shams University, Egypt
Mohamed F. Tolba, Ain Shams University, Egypt

Chapter 56
Research Perspectives on Functional Micro and Nano Scale Coatings: New Advances in Nanocomposite Coatings for Severe Applications..... 1076
Jaime Andrés Pérez Taborda, Universidad Tecnológica de Pereira, Colombia & Instituto de Microelectrónica de Madrid-CSIC, Colombia
Elvis O. López, Centro Brasileiro de Pesquisas Físicas, Brazil

Chapter 57
Optimal Intervention Methods for Markovian Gene Regulatory Networks 1110
Mohammadmahdi Rezaei Yousefi, The Ohio State University, USA

Chapter 58
Human-Data Interaction in Healthcare 1148
Federico Cabitza, Università degli Studi di Milano-Bicocca, Italy
Angela Locoro, Università degli Studi di Milano-Bicocca, Italy

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 59
An Optimized Graph-Based Metagenomic Gene Classification Approach: Metagenomic Gene Analysis 1168
Md Sarwar Kamal, East West University, Bangladesh
Mohammad Ibrahim Khan, Chittagong University of Engineering and Technology, Bangladesh
Kaushik Dev, Chittagong University of Engineering and Technology, Bangladesh
Linkon Chowdhury, Chittagong University of Engineering and Technology, Bangladesh
Nilanjan Dey, Techno India College of Technology, India

Chapter 60
Healthy 3.0: Healthcare Digital Dimensions 1193
Maria Teresa Borges Tiago, University of Azores, Portugal
Flávio Tiago, UAC, Portugal
Francisco Emanuel Batista Amaral, University of Azores, Portugal
Sandra Silva, Unidade de Saúde de Ilha de S. Miguel Azores, Portugal

Chapter 61
Informed Consent in Healthcare: A Study Case of Genetic Services 1222
Bo Yu, George Mason University, USA
Duminda Wijesekera, George Mason University, USA
Paulo Costa, George Mason University, USA

Chapter 62
Towards User Authentication Requirements for Mobile Computing 1254
Yaira K. Rivera Sánchez, University of Connecticut, USA
Steven A. Demurjian, University of Connecticut, USA

Chapter 63
Perceptions of Play: Using Play-Doh to Enhance the Student Experience in Bioscience Higher Education 1283
Gemma Lace-Costigan, University of Salford, UK

Chapter 64
Pre-Implementation Case Studies Evaluating Workflow and Informatics Challenges in Private Primary Care Clinics for Electronic Medical Record Implementation 1297
Calvin Or, The University of Hong Kong, China

Chapter 65
Extraction of Protein Sequence Motif Information using Bio-Inspired Computing 1306
Gowri Rajasekaran, Periyar University, India
Rathipriya R, Periyar University, India

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 66
Classification of Breast Thermograms Using Statistical Moments and Entropy Features with Probabilistic Neural Networks 1328
Natarajan Sriraam, M. S. Ramaiah Institute of Technology, India
Leema Murali, M. S. Ramaiah Institute of Technology, India
Amoolya Girish, M. S. Ramaiah Institute of Technology, India
Manjunath Sirur, M. S. Ramaiah Institute of Technology, India
Sushmitha Srinivas, M. S. Ramaiah Institute of Technology, India
Prabha Ravi, M. S. Ramaiah Institute of Technology, India
B. Venkataraman, Indira Gandhi Centre for Atomic Research, India
M. Menaka, Indira Gandhi Centre for Atomic Research, India
A. Shenbagavalli, National Engineering College, India
Josephine Jeyanathan, Kalasalingam University, India

Chapter 67
HIPAA Security and Privacy Rules Auditing in Extreme Programming Environments 1341
Mahmood Alsaadi, Princess Sumaya University for Technology, Jordan
Malik Qasaimeh, Princess Sumaya University for Technology, Jordan
Sara Tedmori, Princess Sumaya University for Technology, Jordan
Khaled Almakadmeh, Hashemite University, Jordan

Chapter 68
The Cloud Gets Personal: Perspectives on Cloud Computing for Personalized Medicine 1364
Rhoda Joseph, Penn State University, USA
Patrick Brown, Thermo Scientific NanoDrop Products, USA

Chapter 69
Analysis of Healthcare Workflows in Accordance with Access Control Policies 1378
Sandeep Lakaraju, Dakota State University, USA
Dianxiang Xu, Boise State University, USA
Yong Wang, Dakota State University, USA

Chapter 70
Big Data and IoT-Allied Challenges Associated With Healthcare Applications in Smart and Automated Systems 1401
Mamata Rath, C.V. Raman College of Engineering, India

Section 5
Organizational and Social Implications

Chapter 71
Shared Electronic Health Records as Innovation: An Australian Case 1416
Troy Pullen, Queensland State Archives, Australia
Latif Al-Hakim, University of South Australia, Australia

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 72
Electronic Medical Record Implementation Challenges for the National Health System in Greece 1450
Dimitrios G. Katehakis, Institute of Computer Science (ICS), Greece

Chapter 73
Ethics of Electronic Health Record Systems 1467
Brian J. Galli, Long Island University – Post, USA

Chapter 74
Hierarchy Similarity Analyser-An Approach to Securely Share Electronic Health Records 1485
Shalini Bhartiya, IITM, India
Deepti Mehrotra, Amity University, India
Anup Girdhar, Sedulity Groups, India

Chapter 75
Patient Data De-Identification: A Conditional Random-Field-Based Supervised Approach 1502
Shweta Yadav, Indian Institute of Technology Patna, India
Asif Ekbal, Indian Institute of Technology Patna, India
Sriparna Saha, Indian Institute of Technology Patna, India
Parth S Pathak, ezDI, LLC, India
Pushpak Bhattacharyya, Indian Institute of Technology Patna, India

Chapter 76
Mobile Application for Patients' Waiting Time Control and Management of Diagnostic Imaging Examinations 1522
Dimitrios Zarakovitis, University of Peloponnese, Greece
Dimitrios Tsoromokos, University of Peloponnese, Greece
Nikolaos Tsaloukidis, University of Peloponnese, Greece
Athina Lazakidou, University of Peloponnese, Greece

Volume IV

Chapter 77
Ensuring Privacy of Participants Recruited via Social Media: An Australian Retrospective Visualisation and Roadmap 1537
Chandana Unnithan, Torrens University Australia, Australia
Paula M Swatman, Swinburne University, Australia
Jo-Anne Kelder, University of Tasmania, Tasmania

Chapter 78
Implications of Molecular Docking Assay for Bioremediation 1556
Zarrin Basharat, Fatima Jinnah Women University, Pakistan
Monazza Bibi, Fatima Jinnah Women University, Pakistan
Azra Yasmin, Fatima Jinnah Women University, Pakistan

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 79
The Performance of Educational Institutions Through the Electronic Records Management Systems: Factors Influencing Electronic Records Management System Adoption 1578
Muaadh Mukred, National University of Malaysia (UKM), Malaysia
Zawiyah M. Yusof, National University of Malaysia (UKM), Malaysia

Chapter 80
Online Social Exergames for Seniors: A Pillar of Gamification for Clinical Practice..... 1599
Stathis Th. Konstantinidis, NORUT, Norway
Ellen Brox, NORUT, Norway
Per Egil Kummervold, NORUT, Norway
Josef Hallberg, Luleå University of Technology, Sweden
Gunn Evertsen, NORUT, Norway
Johannes Hirche, Luleå University of Technology, Sweden

Chapter 81
Exploring the Role of Service Quality and Knowledge for Mobile Health Services 1632
Nabila Nisha, North South University, Bangladesh
Mehree Iqbal, North South University, Bangladesh
Afrin Rifat, North South University, Bangladesh
Sherina Idrish, North South University, Bangladesh

Chapter 82
Effect of Power and Phase Synchronization in Multi-Trial Speech Imagery..... 1654
Sandhya Chengaiyan, SSN College of Engineering, India
Divya Balathayil, SSN College of Engineering, India
Kavitha Anandan, SSN College of Engineering, India
Christy Bobby Thomas, M.S.Ramaiah University of Applied Sciences, India

Chapter 83
Nursing Students' Perception of Medical Information Protection in Hospitals..... 1674
Hyeon-Cheol Jeong, Sahmyook University, Korea

Chapter 84
Information Security Standards in Healthcare Activities 1683
José Gaivéo, Polytechnic Institute of Setubal, Portugal

Chapter 85
The Politics of Establishing ICT Governance for Large-Scale Healthcare Information Infrastructures 1704
Gro-Hilde Ulriksen, University Hospital North Norway, Norway
Rune Pedersen, University Hospital North Norway, Norway
Gunnar Ellingsen, Arctic University of Norway, Norway

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 86
Inhibitors of Physicians' Use of Mandatory Hospital Information Systems (HIS) 1720
Heiko Gewald, Neu-Ulm University of Applied Sciences, Germany
Corinna Gewald, Neu-Ulm University of Applied Sciences, Germany

Chapter 87
Using Key Performance Indicators to Reduce Perceived Perioperative Complexity and Improve Patient Workflow 1738
Jim Ryan, Auburn University at Montgomery, USA
Barbara Doster, University of Alabama, USA
Sandra Daily, Duke Medical Center, USA
Carmen Lewis, Troy University, USA

Chapter 88
Building Intelligent Systems for Paying Healthcare Providers and Using Social Media to Detect Fraudulent Claims 1758
Jack S. Cook, The College at Brockport, USA
M. Pamela Neely, The College at Brockport, USA

Chapter 89
Forensic Investigation of Digital Crimes in Healthcare Applications 1781
Nourhene Ellouze, University of Carthage, Tunisia
Slim Rekhis, University of Carthage, Tunisia
Noureddine Boudriga, University of Carthage, Tunisia

Section 6
Critical Issues and Challenges

Chapter 90
Bidirectional Spreading Activation Method for Finding Human Diseases Relatedness Using Well-Formed Disease Ontology 1814
Said Fathalla, University of Bonn, Germany & Alexandria University, Egypt
Yaman M. Khalid Kannot, Arab Academy for Science, Egypt

Chapter 91
An Investigation Into the Efficacy of Deep Learning Tools for Big Data Analysis in Health Care 1826
Rojalina Priyadarshini, KIIT University, India
Rabindra K. Barik, KIIT University, India
Chhabi Panigrahi, University of Rajastan, India
Harishchandra Dubey, The University of Texas at Dallas, USA
Brojo Kishore Mishra, C. V. Raman College of Engineering, India

Chapter 92
Real Time Analysis Based on Intelligent Applications of Big Data and IoT in Smart Health Care Systems 1839
Mamata Rath, C.V. Raman College of Engineering, India

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 93
The Impact of Healthcare Information Technology on Patient Outcomes 1858
Edward T. Chen, University of Massachusetts Lowell, USA

Section 7
Emerging Trends

Chapter 94
Role of Bioinformatics in Nanotechnology: An Initiation towards Personalized Medicine..... 1875
Harishchander Anandaram, Sathyabama University, India

Chapter 95
Overview of Cellular Computing-Basic Principles and Applications 1895
Amit Das, University of Kalyani, India
Rakhi Dasgupta, University of Kalyani, India
Angshuman Bagchi, University of Kalyani, India

Chapter 96
From Telecytology to Mobile Cytopathology: Past, Present, and Future..... 1921
Abraham Pouliakis, National and Kapodistrian University of Athens, Greece
Stavros Archondakis, 401 Military Hospital, Greece
Niki Margari, National and Kapodistrian University of Athens, Greece
Petros Karakitsos, National and Kapodistrian University of Athens, Greece

Chapter 97
Examining the Influence of Information Technology on Modern Health Care..... 1943
Edward T. Chen, University of Massachusetts, Lowell, USA

Chapter 98
Early Detection of Dementia: Advances, Challenges, and Future Prospects..... 1963
Stefanos Xefteris, Aristotle University of Thessaloniki, Greece
Evdokimos Konstantinidis, Aristotle University of Thessaloniki, Greece
Antonis S. Billis, Aristotle University of Thessaloniki, Greece
Panagiotis Antoniou, Aristotle University of Thessaloniki, Greece
Charis Styliadis, Aristotle University of Thessaloniki, Greece
Evangelos Paraskevopoulos, Aristotle University of Thessaloniki, Greece
Panagiotis Emmanouil Kartsidis, Aristotle University of Thessaloniki, Greece
Christos A. Frantzidıs, Aristotle University of Thessaloniki, Greece
Panagiotis Bamidis, Aristotle University of Thessaloniki, Greece

Chapter 99
Opportunities and Challenges of Big Data in Healthcare..... 1989
Wafaa Faisal Mukhtar, Sudan University of Science and Technology, Sudan
Eltayeb Salih Abuelyaman, University of Dammam, Saudi Arabia

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 100
Wireless Body Area Network for Healthcare Applications 2002
Danda B. Rawat, Georgia Southern University, USA
Sylvia Bhattacharya, Georgia Southern University, USA

Chapter 101
Information and Communication Technologies in the Healthcare: Future Trends for Project Success 2018
Jorge Gomes, Universidade de Lisboa, Portugal
Mário Romão, Universidade de Lisboa, Portugal

Chapter 102
Robotics in Surgical Techniques Robotics in Surgical Techniques: Present and Future Trends 2033
Vassilia Costarides, National Technical University of Athens, Greece
Apollon Zygomalas, University of Patras, Greece
Kostas Giokas, National Technical University of Athens, Greece
Dimitris Koutsouris, National Technical University of Athens, Greece

Chapter 103
A Next Generation of IT: A Health Record Example with Terminology Binding 2048
Anthony J. Mallia, Edmond Scientific Company, USA

Index xxx

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Preface

Advancements in data science have created opportunities to sort, manage, and analyze large amounts of data more effectively and efficiently. In contemporary society, medical records and patient information have increasingly become more digitized, with patients required to use patient portals to access their medical information. Additionally, doctors, nurses, hospital staff, and specialists rely on electronic documents that are easy to access and share and that can follow patients that require multiple medical experts. Applying new data science technologies to the healthcare industry, which has vast quantities of patient and medical data and is increasingly becoming more data-reliant, is crucial for refining medical practices and patient care as well as maintaining efficient communication between departments and offices.

Staying informed of the most up-to-date research trends and findings is of the utmost importance. That is why IGI Global is pleased to offer this four-volume reference collection that will shed light on practical applications of healthcare analytics for improved patient care, resource allocation, and medical performance, as well as for diagnosing, predicting, and identifying at-risk populations by providing both broad and detailed perspectives on cutting-edge theories and developments. This reference is designed to act as a single reference source on conceptual, methodological, technical, and managerial issues and to provide insight into emerging trends and future opportunities on the connections between data analytics in the field of medicine.

Data Analytics in Medicine: Concepts, Methodologies, Tools, and Applications is organized into seven distinct sections that provide comprehensive coverage of important topics. The sections are:

1. Fundamental Concepts and Theories;
2. Development and Design Methodologies;
3. Tools and Technologies;
4. Utilization and Applications;
5. Organizational and Social Implications;
6. Critical Issues and Challenges; and
7. Emerging Trends.

The following paragraphs provide a summary of what to expect from this invaluable reference tool.

Section 1, "Fundamental Concepts and Theories," serves as a foundation for this extensive reference tool by addressing crucial theories essential to the understanding of data analytics in medicine. Introducing the book is "A Guideline to Use Activity Theory for Collaborative Healthcare Information Systems Design," by Prof. Carolin Durst (University of Bremen, Germany), Prof. Nilmini Wickramasinghe (Epworth HealthCare, Australia & Deakin University, Australia), and Prof. Jana Riechert (University

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of Erlangen-Nuremberg, Germany): a great foundation laying the groundwork for the basic concepts and theories that will be discussed throughout the rest of the book. Section 1 concludes and leads into the following portion of the book with a nice segue chapter, "Population Health Management and the Science of Individuality," by Prof. Anastasius Moumtzoglou (Hellenic Society for Quality and Safety in Healthcare, Greece, & P. & A. Kyriakou Children's Hospital, Ilion, Greece) and Prof. Abraham Pouliakis (2nd Department of Pathology, National and Kapodistrian University of Athens, Athens, Greece).

Section 2, "Development and Design Methodologies," presents in-depth coverage of the conceptual design and architecture of data analytics in medicine. Opening the section is "A Yes/No Answer Generator Based on Sentiment-Word Scores in Biomedical Question Answering," by Prof. Mourad Sarrouti (Laboratory of Computer Science and Modeling, FSDM, Sidi Mohammed Ben Abdellah University, Fez, Morocco) and Prof. Said Ouatik El Alaoui (Laboratory of Computer Science and Modeling, FSDM, Sidi Mohammed Ben Abdellah University, Fez, Morocco). Through case studies, this section lays excellent groundwork for later sections that explore present and future applications for data analytics in medicine. The section concludes with an excellent work by Prof. Basant Tiwari (Hawassa University, Awassa, Ethiopia) and Prof. Vivek Tiwari (IIIT-NR, Naya Raipur, India), "An Intelligent Multi-Objective Framework of Pervasive Information Computing."

Section 3, "Tools and Technologies," presents extensive coverage of the various tools and technologies used in the implementation of data analytics in medicine. The first chapter, "Detecting Human Diseases Relatedness: A Spreading Activation Approach Over Ontologies," by Prof. Said Fathalla (Bonn University, Bonn, Germany & Alexandria University, Alexandria, Egypt), lays a framework for the types of works that can be found in this section. The section concludes with "A Study on Lifetime Enhancement and Reliability in Wearable Wireless Body Area Networks," by Prof. Saranya Vasanthamani (Sri Krishna College of Engineering and Technology, Coimbatore, India). Where Section 3 described specific tools and technologies at the disposal of practitioners, Section 4 describes the use and applications of the tools and frameworks discussed in previous sections.

Section 4, "Utilization and Applications," describes how the broad range of data analytics in medicine efforts has been utilized and offers insight on and important lessons for their applications and impact. The first chapter in this section is "Pluralistic Tendencies in Healing Abantu Illness in the Contemporary South Africa," written by Prof. Kholekile Hazel Ngqila (Walter Sisulu University, South Africa). This section includes the widest range of topics because it describes case studies, research, methodologies, frameworks, architectures, theory, analysis, and guides for implementation. The breadth of topics covered in the section is also reflected in the diversity of its authors, from countries all over the globe. The section concludes with "Big Data and IoT-Allied Challenges Associated With Healthcare Applications in Smart and Automated Systems," by Prof. Mamata Rath (Department of IT, C.V. Raman College of Engineering, Bhubaneswar, India), a great transition chapter into the next section.

Section 5, "Organizational and Social Implications," includes chapters discussing the organizational and social impact of data analytics in medicine. The section opens with "Shared Electronic Health Records as Innovation: An Australian Case," by Prof. Troy Pullen (Queensland State Archives, Australia) and Latif Al-Hakim (University of South Australia, Australia). This section focuses exclusively on how these technologies affect human lives, either through the way they interact with each other or through how they affect behavioral/workplace situations. The section concludes with "Forensic Investigation of Digital Crimes in Healthcare Applications," by Prof. Nourhene Ellouze (University of Carthage, Tunisia), Prof. Slim Rekhis (University of Carthage, Tunisia), and Prof. Noureddine Boudriga (University of Carthage, Tunisia).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 6, "Critical Issues and Challenges," presents coverage of academic and research perspectives on analytical tools and applications for maintaining, monitoring, and storing data, among other functions. The section begins with "Bidirectional Spreading Activation Method for Finding Human Diseases Relatedness Using Well-Formed Disease Ontology," by Prof. Said Fathalla (Enterprise Information Systems [EIS], University of Bonn, Bonn, Germany & Faculty of Science, Alexandria University, Alexandria, Egypt) and Prof. Yaman M. Khalid Kannot (Department of Computer Engineering, Arab Academy for Science, Alexandria, Egypt). Chapters in this section examine theoretical approaches and offer alternatives to crucial questions on the use of data analytics in medicine. The section concludes with "The Impact of Healthcare Information Technology on Patient Outcomes," by Prof. Edward T. Chen (University of Massachusetts Lowell, Lowell, USA).

Section 7, "Emerging Trends," highlights areas for future research within the field of data analytics and their specific application within the medical field and opens with "Role of Bioinformatics in Nanotechnology: An Initiation towards Personalized Medicine," by Prof. Harishchander Anandaram (Sathyabama University, India). This section contains chapters that delve into what might happen in the coming years that can extend the already staggering amount of applications for data analytics in medicine. The final chapter of the book looks at an emerging application within the field in the excellent contribution, "A Next Generation of IT: A Health Record Example with Terminology Binding," by Prof. Anthony J. Mallia (Edmond Scientific Company, North Andover, USA).

Although the primary organization of the contents in this multi-volume work is based on its seven sections, offering a progression of coverage of the important concepts, methodologies, technologies, applications, social issues, and emerging trends, the reader can also identify specific contents by utilizing the extensive indexing system listed at the end of each volume. As a comprehensive collection of research on the latest findings related to using technology to providing various services, *Data Analytics in Medicine: Concepts, Methodologies, Tools, and Applications* provides researchers, doctors, hospital staff, medical administrators, nurses, IT consultants, software developers, medical directors, academicians, students, and all medical practitioners with a complete understanding of the development of applications and concepts in data analytics when applied to the medical field. Given the vast number of issues concerning usage, failure, success, policies, strategies, and applications of data analytics in medicine in countries around the world, *Data Analytics in Medicine: Concepts, Methodologies, Tools, and Applications* addresses the demand for a resource that encompasses the most pertinent research in technologies being employed to globally bolster the knowledge and applications of this trending topic.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 1

Fundamental Concepts and Theories

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 1
A Guideline to Use Activity Theory for Collaborative Healthcare Information Systems Design

Carolin Durst
University of Bremen, Germany

Nilmini Wickramasinghe
Epworth HealthCare, Australia & Deakin University, Australia

Jana Riechert
University of Erlangen-Nuremberg, Germany

ABSTRACT

In healthcare, collaborative systems serve very different stakeholders: researchers, clinicians, nurses, patients, and administrators for instance. In order to design a collaborative healthcare information system that satisfies the stakeholders' needs, all individual requirements have to be mapped into the software. Traditional system design focuses at technical features and oftentimes ignores social requirements like human factors or organizational structures. Instead of integrating existing processes and working habits into a system, a technical-focused design approach tries to force new behaviors. As a result, stakeholders could refuse to use collaborative healthcare information systems. In line with Kuutti (1991), we recommend activity theory as a rich framework to study and design collaborative information systems. We believe that activity theory analysis is particularly useful for healthcare settings where diverse stakeholders pursue very different goals. This book chapter offers a structured approach to analyze collaborative activities and to design the IT-support accordingly.

DOI: 10.4018/978-1-7998-1204-3.ch001

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The development of collaborative healthcare information systems is a very complex but promising undertaking. Numerous collaborative activities can be optimized by information and communication technologies. The increasing digitalization of those activities and processes results in time saving, error reduction and most important it enables the analysis, evaluation and knowledge creation of the collected data. Healthcare information systems can improve and support patient safety and avoid medical errors (Baxter & Rooksby, 2011). To date, all OECD countries recognize the opportunities of those systems especially in the light of access, quality and value facing healthcare delivery (Wickramasinghe & Schaffer, 2010).

However, system failures can lead to damages with serious consequences. A significant example is the failure of the dispatch system of the London Ambulance in 1992. Up to 30 people died because of technical problems. The system was not able to address and support the established practices of dispatch. This example shows that development approaches for collaborative systems in healthcare need to focus on both: technical aspects and work practices (Baxter & Rooksby, 2011).

In order to raise the aforementioned potentials, the major challenge is the design of such collaborative systems in healthcare due to the diverse stakeholders and complex work practices. Traditional system design approaches often target at technologically driven features. This means, instead of existing processes and work habits forming the system, the design approach tries to push new behaviors (Sawyer, 2013). It is widely acknowledged that adopting a socio-technical approach to system development leads to systems that are more acceptable to end users, delivers better value to stakeholders and overall value to the company. Despite this knowledge, there is a lack of theories and methods which consider human, social and organizational factors (Sawyer, 2013; Whitworth & Ahmad, 2014). In order to design effective systems, it is not sufficient to only understand the process in general, but it must also be clear how the various stakeholders interact and what motivates them to strive for the joint objective (Sadeghi, Andreev, Benyoucef, Momtahan, & Kuziemsky, 2014a).

One attempt to narrow the so called socio-technical gap is provided by Kuuttii (1991). He suggested to employ the concept of “activities” - originating from activity theory - as basic unit of analysis for collaborative information systems design. He argues, that the properties of the activity concept correlate well with the characteristics and social requirements of collaborative activities (Kuutti, 1991). Applying activity theory analysis helps to understand the work practices of all stakeholders involved in the activity and how these individuals interact with technological and non-technological artefacts. With creating and analysing an activity system it is possible to obtain a complete overview of an activity. This overview can not only be used to uncover contradictions between components of the activity systems (e.g. a nurse has problems to work with certain tool) but also to have deeper insights into the activity system e.g. the number of stakeholders, the objectives that are motivating the individuals to fulfil an activity, or the number and quality tools used. The results can be used for both analysing an activity and capturing requirements for collaborative healthcare information systems.

The remainder of the book chapter is structured as follows: After presenting the theoretical background of activity theory, we discuss related work on the application of activity theory in healthcare. Based on this foundation, we propose a guideline to use activity theory for the analysis and the design of collaborative healthcare information systems. The book chapter closes with concluding remarks, pointing to areas of our future research.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

Activity Theory

Activity theory has been used in the field of information systems and general HCI research before with promising results (e. g., (Barki, Titah, & Boffo, 2007; Bharosa, Lee, Janssen, & Rao, 2012; Durst & Grottke, 2015; Kamaruddin, Yusop, & Ali, 2011; Riechert, Durst, & Wickramasinghe, 2016; Sadeghi et al., 2014a). The concept, evolved from Russian philosophy, is considered to be a meta-theory or a framework, rather than a predictive theory and aims at providing holistic descriptions for social systems. In this context, an activity is a goal-directed interaction of a *subject* with an *object* mediated by the use of *tools* in order to achieve a certain *outcome* (Leont'ev, 1978; Vygotsky, 1978). These tools can be external such as a scissor or internal such as concepts or heuristics (B. Nardi, 1996).

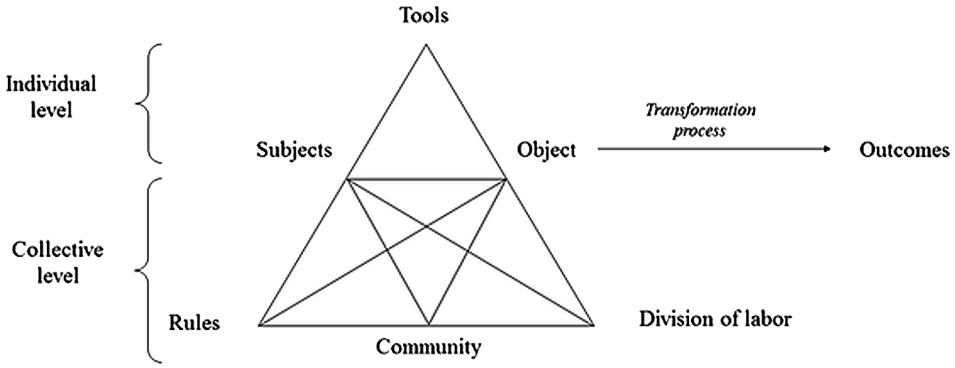
Engeström (1987) extended the original concept of activity to the case of collective activities. He added the element of *community* to the foregoing notion of subject-object interaction in order to enable a three-way interaction between subject, object, and community. Additionally, two types of mediators for each possible interaction between these three elements were defined:

1. *Rules* were introduced to mediate the subject-community interaction.
2. The community-object interaction is mediated by *division of labor* (Engeström, 2000).

The expanded activity system by Engeström can be seen in Figure 1.

According to the principles of activity theory, human activities are always directed towards their objects (Kaptelinin & Nardi, 2012a). As stated by Leont'ev (Leont'ev, 1978), an objectless activity is impossible. "Objects motivate and direct activities, around them activities are coordinated, and in them activities are crystallized when the activities are complete." (Kaptelinin & Nardi, 2012a). Taking on the perspective of a subject, this means that the subject's interactions are organized around objects to achieve an outcome. Analyzing the object of an activity helps to understand activities on an individual and collective level. When defining object, objective, and outcome of an activity it is important to consider the differences between these elements. "Objects have their "objective" meanings, determined by their relationship with other entities existing in the world (including the subject himself or herself)" (Kaptelinin & Nardi, 2012a). The subjects' objectives can be very different, especially in activity systems which

Figure 1. Activity system model of Engeström (2000)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

involve heterogeneous stakeholders. Objectives can not only be physical but also socially or culturally motivated (Kaptelinin & Nardi, 2012a). In a nutshell: Activity theory explains both the 'how' and the 'why' of human action and thus collaboration.

Activity systems are constantly evolving and by nature internally contradictory. Contradictions in activity systems exist because activities are constantly developing and even most well-planned actions involve failures, disruptions, and unexpected innovations (Kaptelinin & Nardi, 2012a). By analyzing the activity system, the underlying contradictions that lead to these failures, disruptions, or innovations may be recognized and offer possibilities for expansive design adjustments (Engeström, Miettinen, & Punamäki, 1999). In activity theory four types or levels of contradictions can be identified (Kaptelinin & Nardi, 2012a).

- **First-Level Contradictions:** Deal with inner-contradictions of each activity system component: subject, object, community, tools, rules, and division of labor. For example, a physician chooses a more affordable medication over the best available medication that is more expensive (Kaptelinin & Nardi, 2012a).
- **Second-Level Contradictions:** Occur between the components of an activity system. For example, if a certain type of medical treatment is unsuitable for certain patients (Kaptelinin & Nardi, 2012a).
- **Third-Level Contradictions:** Are potential problems emerging between the existing and a more advanced form of the activity system (Kaptelinin, Nardi, & Macaulay, 1999). For example, stakeholders involved in a certain activity could refuse to adapt to new IT-supported work practices.
- **Fourth-Level Contradictions:** Occur between different activity systems that are involved in the achievement of a joint outcome. For example, a positive effect of surgery can be undermined by an improper follow-up rehabilitation (Kaptelinin & Nardi, 2012a).

Activity theory is regarded as a powerful lens for studying complex social systems. It analyzes existing relationships among different elements mediated by tools, rules and division of labor. In addition, activity theory considers different user perspectives of the subjects, too (Greenhow & Belbas, 2007). Even though it is characterized as an analytical framework, activity theory does not offer ready-made techniques and procedures for research (Rogers, 2008). It is considered as an evolving framework that only contains general guidelines and therefore must be further developed (Greenhow & Belbas, 2007).

Activity-Oriented Design Method

In literature, we identified two different methods of executing activity theory analysis: The activity-oriented design method (AODM) by Mwanza (2011) and the activity checklist by Kaptelinin and Nardi (Kaptelinin et al., 1999). The activity checklist offers important aspects for system evaluation and system design. Amongst others, it includes aspects such as "People who use the target technology" or "Criteria for success or failure of achieving target goals". However, there are no specific recommendations of how to gather the needed information (Kaptelinin et al., 1999). The AODM by Mwanza (2011) offers a series of well-structured procedures and stages in order to execute the activity theory analysis (Greenhow & Belbas, 2007). In consequence, we adopt the AODM by Mwanza (2011) for the analysis and the design of collaborative healthcare information systems.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The AODM was developed to identify contradictions in activity systems (Greenhow & Belbas, 2007). According to Engeström, contradictions are seen as means of expanding and transforming the system and developing human practices. In other words, uncovering contradictions is necessary in order to improve system design (Greenhow & Belbas, 2007). In addition, design insights for further studies and refinement can be derived (Greenhow & Belbas, 2007). The AODM offers four methodological tools to support data gathering, analysis and communication of design insights. These tools are divided into four stages (Sadeghi et al., 2014a):

Stage 1: Defining Activity System (Eight Step Model)

The eight step model of stage 1 puts theory into practice by defining and interpreting the activity system. It is a requirements capture methodology grounded in activity theory (D. Mwanza & Engeström, 2003). Table 1 shows the eight step model with the corresponding questions to ask.

Stage 2: Activity Notation

Stage 2 reduces complexity of the examined activity system by decomposing or breaking down the complex activity system into manageable sub-activities. Three rules-of-thumb, proposed by Mwanza (2011), should be considered:

1. The actor is represented by the subject or community component of the activity system;
2. Tools, rules, or division of labor function as mediators; and
3. The activity focuses on an object. As a result of this stage, a triangle for each sub-activity is created (Sadeghi et al., 2014a).

Table 1. The eight step model

Step	Identification of	Question to Ask
1	Activity of interest	What sort of activity am I interested in?
2	Objective	Why is the activity taking place?
3	Subjects	Who is involved in carrying out this activity?
4	Tools	By what means are the subjects performing this activity?
5	Rules and Regulations	Are there any cultural norms, rules or regulations governing the performance of activity?
6	Division of Labor	Who is responsible for what, when carrying out activity and how are the roles organized?
7	Community	What is the environment in which this activity is carried out?
8	Outcome	What is the desired outcome from carrying out this activity?

Mwanza-Simwami, 2011.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Stage 3: Research Questions

At stage 3, a research question for each sub-activity is generated. Here, either generic questions such as "What tools are used by the subjects to achieve their objects and how do they do that?" or questions specific to the context such as "How do whiteboards help the clinical manager to facilitate the safe and effective transition of patients towards discharge?" (Sadeghi et al., 2014a) are asked.

Stage 4: Data Collection and Technique of Mapping Operational Processes

Based on these questions data is collected and analyzed in stage 4. Stage 4 communicates research insight by modelling inter-relations of operational processes and by modelling study findings (e.g., contradictions identified in the analysis of human practices).

The intended result of AODM is to uncover contradictions. In a next step, these contradictions serve as a starting-point to improve the system design. In order to solve the contradiction, elements of the activity system need to be adapted or changed. For instance, the division of labor is revised, new rules are introduced or tools need to be adapted. Another way of adapting the elements of an activity system is to support them with IT. Examples of IT support are the automation of routines supported by tools and providing data about the object or make the object manipulable. Linking work tasks of several people in order to create a community is another example of IT support (B. Nardi, 1996).

Applications of Activity Theory in Healthcare

Since Kaptelinin and Nardi's work "Synthesis Lectures on Human-Centered Informatics: Activity Theory in HCI: Fundamentals and Reflections" (Kaptelinin & Nardi, 2012b), activity theory analysis is widely accepted in information systems research. A plentiful amount of studies was published in the last years. Examples of these studies are "An activity-theory analysis of corporate wikis" (Hasan & Pfaff, 2012), "Using Activity Theory to Model Context Awareness" (Cassens & Kofod-Petersen, 2006), and "Demonstrations of the Activity Theory Framework for Research in IS" (Crawford & Hasan, 2006). So far, the largest field of application is education. Publications in this field are characterized by their attempt to improve education services, for example "Using activity theory and its principle of contradictions to guide research in educational technology" (Murphy & Rodriguez-manzanares, 2008).

In the context of healthcare, however, activity theory is not widely used and we identified only a couple of authors who applied the concept of activity theory. Engeström's research group (Engeström, 2000; Hasu & Engeström, 2000) belong to the first researchers that applied activity theory analysis to traditional work systems in healthcare. After the digitalization had transformed the traditional healthcare processes, Er and Kay (Er & Kay, 2005), Zao et al. (Zao et al., 2008), Bardram and Doryab (Bardram & Doryab, 2011) and (Sadeghi et al., 2014a) revisited this domain. Er and Kay (Er & Kay, 2005) studied technology issues of a mobile hospital ward system through the lens of activity theory. The authors conducted several semi-structured interviews and observed doctors while conducting their ward rounds and mapped their observations into an activity system. Activity theory was found to be a useful approach to identify additional issues. Zao et al. (Zao et al., 2008) also applied activity theory to a mobile healthcare solution. But in contrast to Er and Kay (Er & Kay, 2005) the authors used the theory to design a new application. In their article, they compared activities to features of the desired mobile solution. This approach was found to be very helpful to structure the design process of their mobile solution to support the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

elderly, but according to the basic idea of Engeström, activities are more than just "features". Bardram and Doryab (Bardram & Doryab, 2011) and (Sadeghi et al., 2014a) applied activity theory successfully to the newly emerging socio-technical systems and were able to provide insights for the design of work environments. While Bardram and Doryab (Bardram & Doryab, 2011) used activity theory to describe processes of a hospital in a very detailed manner, Sadeghi et al. (Sadeghi et al., 2014a) applied the AODM to analyze the patient discharge process and identified several contradictions. Both approaches were descriptive and did not aim at deriving design recommendations based on the uncovered contradictions (Sadeghi et al., 2014a).

Riechert et al. (Riechert et al., 2016) also applied AODM to analyze the chemotherapy administration process and derived several contradictions as well as design recommendations. Building prior research and the lessons learned from the chemotherapy case, the authors modified and enhanced the initial AODM. The following section presents the derived learnings by means of a guideline to use activity theory for the analysis and design of collaborative healthcare information systems.

GUIDELINE TO USE ACTIVITY THEORY FOR COLLABORATIVE HEALTHCARE INFORMATION SYSTEM DESIGN

Based on the experiences collected during the chemotherapy case study a guideline for using activity theory analysis was developed by extending the AODM. The guideline is represented in Table 2. Literature-based recommendations are also included.

Finally, it is important to mention that both executing the chemotherapy case study and analyzing literature made obvious that the stages of the AODM cannot be applied iteratively. Insights gathered in one stage often lead to changes in another stages, mostly in the eight-step model. Therefore, the application of the proposed guideline is rather cyclic.

DISCUSSION AND CONCLUSION

The successful development of collaborative information systems in healthcare faces several challenges: The work environment in healthcare is complex and dynamic. Health professionals work with multiple technologies, paper, devices, other people, and patients. Many technologies and tools might be used at the same time. Furthermore, it still is a challenge to produce an integrated electronic record. Even though electronic records are already widely used, there are still problems such as ensuring the correct people have access to records or the information is correct and meaningful. In addition, health professionals still like to work with paper because historical information and other notes can be captured and recognized easily.

Table 2. Guideline for using activity theory analysis based on the AODM

Stage	Recommendations
Stage 0: Preparatory Stage	Collect the data that is needed for the following stages (Greenhow & Belbas, 2007) if scope requires or much data are available

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Continued

Stage		Recommendations
Stage 1: Defining activity system – eight- step model	1.1 General	• Use information from people involved or experienced with the activity of interest rather than literature analysis because then the collected information is more significant • Consider that other people might have no knowledge about activity theory analysis • Start with capturing the subjects as this is the easiest start for someone who has little or no knowledge about activity theory analysis • Start with defining subject or outcome because these are the most obvious o Because the outcome might be defined in the strategic management or o Easier to capture in comparison to other elements of the activity system o Capturing the subjects first might be the easiest for someone who has little or no knowledge about activity theory analysis • In general, capture the more obvious elements first • Develop an activity system for each subject in order to be able to consider the different perspectives of these subjects (Greenhow & Belbas, 2007) if scope requires
	1.2 Activity of interest	• First, define the scope of the activity by selecting a specific activity or sub-activity that is further analyzed o Consider that this decision has a high influence on the complexity of the activity theory analysis o Keep in mind time, costs, and quality requirements of the project o Prefer to develop several activity systems instead of one complex one • Double check if the analyzed point of view is correct
	1.3 Object(ive)	• Differentiate between object, objective, and outcome at once • Object: material or mental and can be modified by a subject according to the activities objective (Bedny, 2014) • Objective: aim of the activity e.g. providing better healthcare. Ask "Why is the activity taking place?", is the direct target or aim of activity • Outcome: the activity's result e.g. improved health
	1.4 Subjects	• List all participants of the activity • Consider how subjects are defined in activity theory: Subjects can be individual humans, animals, teams, and organizations (Kaptelinin & Nardi, 2012b) • Capture subjects and their division of labor simultaneously in order to save time (only feasible in dialogue) • Consider that the roles of subject and object can be switched during the execution of an activity and decide for the appropriate point of view
	1.5 Tools	• For initial analysis, take all tools (electronic, paper-based, communications tools etc.) into consideration to get a rich overview • Especially for projects in information technology: all aspects of activity need to be considered e.g. not only software tools but also cognitive tools. Do not only focus on the software in order to ensure completeness. • If the scope of the activity system is adapted, it may would make sense to keep only specific tools or summarize likewise tools into groups • divide artefacts or tools into groups such as social, assignments, cognitive, and technological (Greenhow & Belbas, 2007) if scope requires
	1.6 Rules and regulations	Consider that this section can get very complex depending on the scope of activity (e.g. if code of conduct or corporate guidelines are included)
	1.7 Division of labor	• Describe responsibilities of all subjects • Capture subjects and their division of labor simultaneously in order to save time (only feasible in dialogue) • Include additional roles such as external roles if the scope requires
	1.8 Community	• Ask "Who has an influence on the object/objective/outcome?" e.g. all employees of a department in an organization because only they have access to comment ideas • If it is not clear from the beginning, take a high scope e.g. all employees and break down later
	1.9 Outcome	• Differentiate between object, objective, and outcome at once • Can be seen as overall (or strategic) goal such as increasing flexibility or reducing healthcare costs

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Continued

Stage	Recommendations
Stage 2: Activity notation	• Consider Mwanza's three rules-of-thumb: (1) the actor is represented by the subject or community component of the activity system; (2) tools, rules, or division of labor function as mediators and (3) the activity needs to focus on an object (Daisy Mwanza, 2002) • Double check if the captured activity system is correct because there might have been confusions during the interview
Stage 3: Research questions	• First, refer all research questions to the objective of the activity in order to remain in scope • Second, add additional research questions if necessary for the specific analysis
Stage 4: Data collection and technique of mapping operational processes	• Use the research of the sub-triangles e.g. subject-tools-object to find contradictions • Use the levels of contradictions in order to capture them completely • Define the activity hierarchy in order to get a more valuable overview of the activity • Use dialogue to capture the information because this proved as most efficiently as discussions uncovered many aspects of the activity system

We found, that activity theory analysis is a useful lens to analyse collaborative activities in healthcare settings and to design the IT support accordingly. So far, activity theory analysis was criticized for its descriptive character without delivering predictions for future development (Nardi 1996 cited in Nathanael et al. 2002). This is true but activity theory analysis offers a high value for system analysis and offers a good starting point for system or design improvements. The chemotherapy case study proved that the uncovered contradictions of an activity system can lead to activity system recommendations by changing or adopting the activity elements.

Capturing the elements of the activity system and the sub-systems is critical for the whole analysis, influences all upcoming steps and decides on success or failure of the analysis (Greenhow & Belbas, 2007). It should be done through the lens of the overall objective of the activity system. To date, there is no clear method offered by the AODM, which helps researcher to define this scope. Existing literature does not differentiate clearly between the activity system elements *object*, *objective* and *outcome*. Some even use object and objective synonymously (Greenhow & Belbas, 2007; Sadeghi, Andreev, Benyoucef, Momtahan, & Kuziemsky, 2014b). In order to find contradictions and to derive design recommendations we found, that this differentiation is very crucial. Our guideline builds on the existing AODM. We added our lessons learned while applying the AODM in terms of hints or additional questions and findings from other literature. We want to encourage researchers to take up the identified suggestions, apply and improve them in order to establish activity theory as a solid foundation for collaborative healthcare information system design.

REFERENCES

Bardram, J., & Doryab, A. (2011). Activity Analysis - Applying Activity Theory to Analyze Complex Work in Hospitals. *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11,* Hangzhou (pp. 455–464). 10.1145/1958824.1958895

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Barki, H., Titah, R., & Boffo, C. (2007). Information system use-related activity: An expanded behavioral conceptualization of individual-level information system use. *Information Systems Research*, *18*(2), 173–192. doi:10.1287/isre.1070.0122

Baxter, G., & Rooksby, J. (2011). *Socio Technical Systems Engineering Handbook. Urban Systems*. St. Andrews: University of St. Andrews. Doi:10.1016/0147-8001(77)90017-1

Bedny, G. Z. (2014). *Application of Systemic-Structural Activity Theory to Design and Training*. Boca Raton: CRC Press. Retrieved from http://site.ebrary.com/lib/erlangen/detail.action?docID=11007138

Bharosa, N., Lee, J., Janssen, M., & Rao, H. R. (2012). An activity theory analysis of boundary objects in cross-border information systems development for disaster management. *Security Informatics*, *1*(1), 15. doi:10.1186/2190-8532-1-15

Cassens, J., & Kofod-Petersen, A. (2006). Using Activity Theory to Model Context Awareness : A Qualitative Case Study. *Proceedings of the Nineteenth International Florida Artificial Intelligence Research Society Conference (FLAIRS)*.

Crawford, K., & Hasan, H. M. (2006). *Demonstrations of the Activity Theory Framework for Research in IS. Demonstrations of the Activity Theory Framework for Research in IS*, *13*(2), 49–68.

Durst, C., & Grottke, M. (2015). An Empirical Analysis of System-generated Data in Location-based Crowdsourcing. *Proceedings of the 12th International Conference on Wirtschaftsinformatik* (pp. 933–947).

Engeström, Y. (1987). *Learning by Expanding An Activity-Theoretical Approach to Developmental Research. Helsinki: Orienta-Konsultit Oy (1. Edition.)*. Helsinki: Orienta-Konsultit Oy.

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, *43*(7), 960–974. doi:10.1080/001401300409143 PMID:10929830

Engeström, Y., Miettinen, R., & Punamäki, R.-L. (1999). *Perspectives on activity theory*. Cambridge University Press. doi:10.1017/CBO9780511812774

Er, M., & Kay, R. (2005). Mobile technology adoption for mobile information systems: an activity theory perspective. *Proceedings of the International Conference on Mobile Business (ICMB'05)* (pp. 322 – 325). 10.1109/ICMB.2005.70

Greenhow, C., & Belbas, B. (2007). Using activity-oriented design methods to study collaborative knowledge-building in e-learning courses within higher education. *International Journal of Computer-Supported Collaborative Learning*, *2*(4), 363–391. doi:10.100711412-007-9023-3

Hasan, H., & Pfaff, C. C. (2012). An activity-theory analysis of corporate wikis. *Information Technology & People*, *25*(4), 423–437. doi:10.1108/09593841211278802

Hasu, M., & Engeström, Y. (2000). Measurement in action: An activity-theoretical perspective on producer-user interaction. *International Journal of Human-Computer Studies*, *53*(1), 61–89. doi:10.1006/ijhc.2000.0375

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kamaruddin, K., Yusop, N., & Ali, M. (2011). Using activity theory in analyzing requirements for mobile phone application. *Proceedings of the 5th Malaysian Conference in Software Engineering (MySEC)* (pp. 7–13). 10.1109/MySEC.2011.6140636

Kaptelinin, V., & Nardi, B. (2012a). *Activity Theory in HCI. Fundamentals and Reflections. Bulletin de correspondance hellénique*. San Rafael, CA, USA: Morgan & Claypool Publishers; doi:10.3406/bch.1986.1808

Kaptelinin, V., & Nardi, B. (2012b). Basic concepts and principles of activity theory. In *Activity Theory in HCI* (pp. 1–105). Fundamentals and Reflections.

Kaptelinin, V., Nardi, B., & Macaulay, C. (1999). Methods & tools: The activity checklist: a tool for representing the "space" of context. *Interaction*, *6*(4), 27–39. doi:10.1145/306412.306431

Kuutti, K. (1991). The concept of activity as a basic unit of analysis for CSCW research. *Proceedings of the Second Conference on European Conference on Computer-Supported Cooperative Work ECSCW'91* (pp. 249–264). 10.1007/978-94-011-3506-1_19

Leont'ev, A. N. (1978). *Activity, Consciousness, and Personality*. Englewood Cliffs, NJ: Prentice Hall.

Murphy, E., & Rodriguez-manzanares, M. a. (2008). Using activity theory and its principle of contradictions to guide research in educational technology. *Australian Journal of Educational Techonologyu*, *24*(4), 442–457. Retrieved from http://www.ascilite.org.au/ajet/ajet24/murphy.pdf

Mwanza, D. (2002). Conceptualising work activity for CAL systems design. *Journal of Computer Assisted Learning, 18*, 84–92. doi:10.1046/j.0266-4909.2001.00214.x

Mwanza, D., & Engeström, Y. (2003). Pedagogical Adeptness in the Design of E-learning Environments: Experiences from the Lab@Future Project. *Proceedings of the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education,* Helsinki, Finland (pp. 1344–1347). Retrieved from http://oro.open.ac.uk/11794/

Mwanza-Simwami, D. (2011). AODM as a framework and model for characterising learner experiences with technology. *Journal of E-Learning and Knowledge Society*, *7*(3), 75–85.

Nardi, B. (1996). Activity theory and human-computer interaction. In B. Nardi (Ed.), *Context and consciousness: activity theory and human-computer interaction* (pp. 7–16). Cambridge, London: MIT Press.

Nardi, B. A. (1996). *Context and Consciousness: Activity Theory and Human Computer Interaction. Camebridge*. MIT Press.

Nathanael, D., Marmaras, N., Papantoniou, B., & Zarboutis, N. (2002). Socio-technical Systems Analysis: Which approach should be followed? [Catania.]. *Proceedings of ECCE '11*, 11–13.

Riechert, J., Durst, C., & Wickramasinghe, N. (2016). The application of activity theory to explain collaborative technology use in healthcare: The case of a chemotherapy ordering system. *Proceeding of the 49th Hawaii International Conference on System Sciences (HICSS)*. 10.1109/HICSS.2016.126

Rogers, Y. (2008). 57 Varieties of Activity Theory. *Interacting with Computers*, *20*(2), 247–250. doi:10.1016/j.intcom.2007.07.004

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sadeghi, P., Andreev, P., Benyoucef, M., Momtahan, K., & Kuziemsky, C. (2014a). Activity Theory driven system analysis of complex healthcare processes. *Proceedings of the European Conference on Information Systems (ECIS) 2014*. Tel Aviv: AIS Electronic Library.

Sadeghi, P., Andreev, P., Benyoucef, M., Momtahan, K., & Kuziemsky, C. (2014b). Activity Theory Driven System Analysis of Complex Healthcare Processes.

Sawyer, S. (2013). Sociotechnical approaches to the study of Information Systems. CRC Handbook of Computing, 1–39.

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes* Cambridge: Harvard University Press.

Whitworth, B., & Ahmad, A. (2014). Socio-technical system design. In M. Soegaard & R. F. Dam (Eds.), *The Encyclopedia of Human-Computer Interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation.

Wickramasinghe, N., & Schaffer, J. L. (2010). *Realizing Value Driven e-Health Solutions*. Washington.

Zao, J. K., Fan, S. C., Wen, M. H., Hsu, C. T., Hung, C. H., Hsu, S. H., & Chuang, M. C. (2008). Activity-oriented design of Health Pal: A smart phone for elders' healthcare support. *EURASIP Journal on Wireless Communications and Networking*, *2008*, 1–27. doi:10.1155/2008/582194

This research was previously published in the Handbook of Research on Healthcare Administration and Management edited by Nilmini Wickramasinghe; pages 616-626; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 2
A Review of Type 1 Diabetes (T1D): To Assist Patients' and Carers' Understanding of This Condition

Eileen O'Donnell
Technological University Dublin, Ireland

Liam O'Donnell
Technological University Dublin, Ireland

ABSTRACT

The diagnosis of Type 1 Diabetes (T1D) will come as an unwelcome surprise to most people. Within a short period of time, the person will have to come to understand and manage this chronic illness. The terminology associated with the T1D condition will also be totally new to the person: diabetes mellitus, pancreas, hyperglycaemia (hyper), hypoglycaemia (hypo), bolus (fast acting insulin), basal (slow acting insulin), ketones and blood glucose levels. The purpose of this article is to assist newly diagnosed patients' understanding of T1D, people who are already living with T1D, carers of people with T1D, partners and family members of someone with T1D, work colleagues, and friends who participate in the same sporting activities or go on holiday with a person who has T1D. In addition, this article reviews how people living with T1D can still enjoy exercise and maintain the best quality of life possible; whilst controlling the blood glucose levels in their body for the rest of their lives to prevent the onset of complications associated with diabetes.

INTRODUCTION

The diagnosis of T1D at any age can be quite shocking. Particularly to an individual or parent who has had no close previous contact with anyone with the condition. "85% of newly-diagnosed patients have no family history of the disease" (Atkinson & Eisenbarth, 2001, p. 224). The diagnosis of T1D, a debilitating autoimmune disease (Wen et al., 2008), has increased substantially over the last thirty years

DOI: 10.4018/978-1-7998-1204-3.ch002

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(Atkinson & Eisenbarth, 2001), and continues to increase (Ang et al., 2017; O'Hara et al., 2017), to the extent that it is now considered to be a global epidemic (Ahmed et al., 2018; Noor, Khan, & Ahmad, 2017; O'Loughlin, McIntosh, Dinneen, & O'Brien, 2010).

T1D is a life-long medical condition requiring frequent monitoring of blood glucose levels throughout the day, constant vigilance of every morsel the patient consumes and other lifestyle choices are necessary to effectively manage this condition (Dinneen, 2010). In addition, the patient or carer must have an appreciation of the effect that personal exercise choices may have on blood glucose levels and how to manage the precise dose of insulin administered to cover for the possible reduction in blood glucose levels experienced during exercise.

Even if a person with Type 1 Diabetes takes great care and manages to keep their blood glucose levels as close to normal as possible, there is still no guarantee that they will not suffer from complications due to diabetes later on in life. Basically, the person has the power to control/modify all of the risk factors through tight monitoring of blood glucose levels, diet, exercise, life style choices, etc. The only factor that is non-modifiable is the fact that a person diagnosed with T1D has it for life, everything else that effects blood glucose levels are modifiable. The pancreas of a person with Type 1 Diabetes no longer produces insulin or produces insufficient doses of insulin to control the amount of glucose in the blood. Therefore, the challenge is to balance insulin doses and everything else (food and drink intake, exercise, sleep, stress, illness, and so forth) to maintain blood glucose levels as near to normal as possible.

T1D is challenging to manage as so many different variables influence the level of glucose in one's blood at any single point in time. "Crucial to achieving glycaemic control is an understanding of the effect of diet, physical activity, and insulin on blood glucose levels" (Jackson et al., 2015, p. 1959). The administration of the most suitable dose of insulin to cover the amount of carbohydrates consumed and the level of activity in which the person engages is paramount to achieving good glycaemic (blood glucose level) control in order to avoid other medical complications in the future. The administration of the most suitable dose is also crucial to ensure that the person who has T1D does not experience a severe hypoglycaemia episode (hypo).

"Type 1 diabetes is a chronic disease in which individuals are unable to produce insulin, which is essential for regulating blood sugar" (Byrne et al., 2012, p. 3), that is characterised by hyperglycaemia and the harm it may cause to the body (Fowler, 2008). Hyperglycaemia means that there is too much glucose in the person's blood stream, which can, if left untreated, lead to other medical complications in later life. Hyperglycaemia can also lead to Diabetic Ketoacidosis (DKA) in the short term. DKA is a life-threatening condition that may affect people who are diagnosed with Type 1 or Type 2 diabetes. Alternatively, a person may present at a Hospital's Accident and Emergency Department with DKA, this being the first sign that they are in fact suffering from diabetes. DKA happens when the levels of glucose (sugar) in the blood rise to dangerous levels because the pancreas is unable to produce sufficient insulin to regulate the level of glucose in the blood. Therefore, the body cannot access glucose from the blood for energy, tissue production and repair, instead the body starts to breakdown stores of fat that are present in the body. This process leads to the production of ketones. The presence of ketones can turn the blood acidic. Acidic blood can make a person feel very unwell, and subsequently if untreated can lead to DKA, a diabetic coma or even death.

The background section discusses the worldwide incidence of T1D and some of the symptoms experienced during the sudden onset of T1D. The section on the diagnosis of T1D includes brief explanations of the following: Self-Monitoring of Blood Glucose (SMBG); blood glucose levels; HbA1c; the use of insulin to control the blood glucose level; acceptable blood glucose level; unique treatment plan

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

required for each patient; pancreas; and the honeymoon phase. The next section covers some factors which may influence the onset of T1D: genetics; environmental; deficiency in Vitamin D; stress; and the use of statins. The following section introduces the equipment, disposables, medication, and the monetary costs required to manage T1D. The responsible administration of insulin is explained in the next section, followed by sections on hypers and hypos. The subsequent section includes some recommendations for the successful management of T1D. Succeeded by some tips for patients with T1D. Then there is a section that briefly mentions medical complications association with diabetes, foot ulceration and nephropathy. The contraindications of administering insulin are noted in the next section. Followed by a section on the positive impact of a diagnosis of T1D, several options worthy of further research, the conclusion, and references.

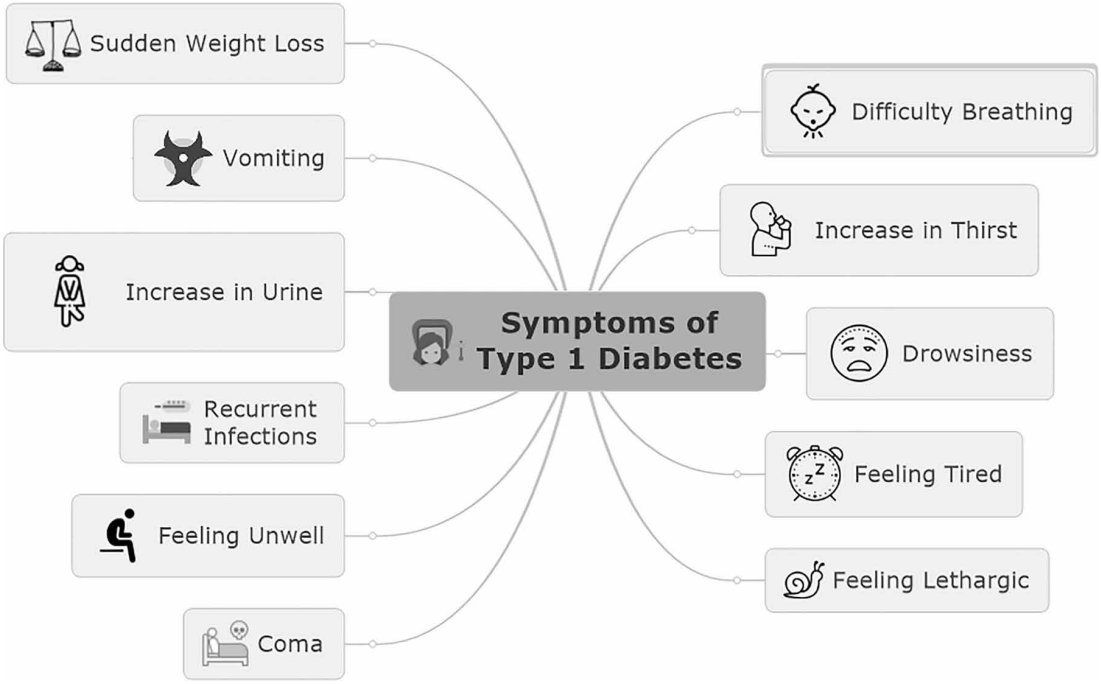
BACKGROUND

Initially, a diagnosis of T1D may sound very daunting and the requirement to self-manage this complicated condition may seem insurmountable or at least very challenging. "While there has been a focused shift away from the traditional medical models of practice where the clinician is perceived as the expert who holds a position of authority over the patient, some clinicians still struggle with the shift to shared control" (Hurley et al., 2017, p. 1958). A person with T1D must take personal responsibility for injecting themselves with insulin several times a day and also for deciding how many units of insulin to inject at each meal because they may only have a consultation with the clinician once or twice a year.

Type 1 Diabetes Mellitus (T1DM), also known as Type 1 Diabetes (T1D) accounts for 5-10% of the total number of cases of Diabetes Mellitus worldwide (Voltarelli et al., 2007), or approximately 20 million people globally (Thrower & Bingley, 2014). "The term diabetes mellitus describes a metabolic disorder of multiple aetiology characterized by chronic hyperglycaemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both" (Alberti et al., 1998, p. 2). "Hyperglycemia is associated with increased mortality in critically ill patients" (Griesdale et al., 2008, p. 821). "Type 1 diabetes has historically been considered a disorder predominately of children and young adults (the disease is commonly referred to as juvenile diabetes because it has peak expression between ages 10-14 years)" (Atkinson & Eisenbarth, 2001, p. 223). The onset of T1D usually occurs before the age of thirty years (Hex, Bartlett, Wright, Taylor, & Varley, 2012). Young adults with T1D struggle to manage blood glucose levels leading to poor glycaemic control and suboptimal clinical outcomes (Byrne et al., 2017; O'Hara et al., 2016), therefore, "effective interventions to improve self-management and outcomes for young adults with T1DM are needed" (Byrne et al., 2017, p. 2). "The precise incidence of new-onset type 1 diabetes in those over 20 years of age is unknown" (Chiang, Kirkman, & Laffel, 2014, p. 2034). "Half of people with type 1 diabetes are diagnosed as adults" (Gilbert, Millard, & Flynn, 2012, p. 2). "The clinical diagnosis is often prompted by symptoms such as increased thirst and urine volume, recurrent infections, unexplained weight loss and, in severe cases, drowsiness and coma" (Alberti et al., 1998, p. 4). Symptoms can also include difficulty breathing, vomiting and a general sense of feeling unwell. Please see Figure 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Some of the symptoms of type 1 diabetes



DIAGNOSIS OF T1D

The determination of the presence of T1D in a patient can be assessed through various different medical tests: (i) Fasting blood glucose level, (ii) Oral Glucose Tolerance Test (OGTT), (iii) HbA_{1c} test which provides an average of glycated haemoglobin over approximately the last two to three months, (iv) a random blood glucose level greater than or equal to 11.1mmol/L (200mg/dL).

Self-Monitoring of Blood Glucose (SMBG)

A person who has recently been diagnosed with T1D can broaden and enhance their understanding of this condition by (i) following the advice of the medical professionals on the assigned diabetes care team, (ii) getting the opportunity to discuss the experiences of others who have lived with diabetes and how they have managed their condition, and (iii) reading books, journal articles and online forums. One of the main responsibilities of a person diagnosed with T1D is to avoid experiencing hypos and hypers through regular Self-Monitoring of Blood Glucose (SMBG) (Deiss et al., 2006; Garg, Shah, Akturk, Beatson, & Snell-Bergeon, 2017), and by improving their understanding of the condition by engaging with Structured Education Programmes (SEPs) (Lawton & Rankin, 2010). Unfortunately, Diabetes Self-Management Education (DSME) does not receive the same level of funding as "drugs and devices (which are associated with financial gain)" (Hurley et al., 2017, p. 1958). Achieving near to normal blood glucose levels is challenging (Russell et al., 2014), but very important, as hyperglycaemia is associated with many negative cardiovascular and immunological effects (Egan & Dinneen, 2016). Cardiovascular disease affects the heart, blood vessels and arteries.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Blood Glucose Levels

T1D can be diagnosed from a fasting blood glucose level greater than or equal to 7 mmol/L (126 mg/dL) (Atkinson, Eisenbarth, & Michels, 2014; Chiang et al., 2014; Daneman, 2006; Egan & Dinneen, 2014). "The cut-off point used to define a normal plasma glucose differs between the USA (where the American Diabetes Association (ADA) recommends a concentration of 5.6) mmol/L and the rest of the world (where the World Health Organisation (WHO) recommends 6.1 mmol/L)." (Dinneen, 2010, p. 589). OGTT is the abbreviation for the Oral Glucose Tolerance Test. A Glucose Tolerance Test (GTT) result of 11.01 mmol/L (200mg/dL) or greater (Atkinson et al., 2014; Chiang et al., 2014; Egan & Dinneen, 2014) signifies the presence of diabetes. "The diagnostic cut off points of 7.0 mmol/L (fasting) and 11.1 mmol/L (OGTT 2-hour value) are based on the concentrations at which retinopathy begins to appear in a population" (Egan & Dinneen, 2014, p. 680). Retinopathy is the medical term for damage to blood vessels in the eyes due to high blood sugar levels. The diagnostic cut off point of haemoglobin A1c > 7.0% is associated with patients who are at increased risk of developing Cardiovascular Autonomic Neuropathy (CAN) (Spallone et al., 2011). Neuropathy is the medical term for damage to peripheral nerves resulting in weakness, numbness, and pain. The risk of developing retinopathy and other complications associated with diabetes depends on the duration and severity of the hyperglycaemic episodes experienced (Fowler, 2008). Malone, Pavan, Morrison, & Cuthbertson (2001) suggest that dilated eye examinations that are performed to detect retinopathy should be conducted within the first year after the onset of T1D; as their study found that 20% of the volunteers who lived with diabetes for 1 year tested positive for retinopathy.

HbA_{1c}

People with poor glycaemic control have haemoglobin HbA_{1c} levels in excess of 7% (Davis, 2016). "The American Diabetes Association currently recommends an A1c goal of less than 7.0%, while other groups such as the American Association of Clinical Endocrinologists recommend a goal of less than 6.5%" (Davis, 2016, p. 4). A HbA_{1c} test of 6.5% or higher indicates the presence of diabetes (Egan & Dinneen, 2014). A HbA_{1c} test provides an average of glycated haemoglobin (Byrne et al., 2012) in the blood over the past two to three months (Atkinson et al., 2014). In Dinneen et al.'s (2013) study one of the inclusion criteria to participate in the study was to have "an HbA_{1c} level below 13 percent at recruitment" (Dinneen et al., 2013, p. 30). Dinneen et al. (2013) in their study "Based on pre-study audit data from Irish centres estimated that approximately 20 percent of participants would have a baseline HbA_{1c} below 7.5 percent (59 mmol/mol)" (Dinneen et al., 2013, p. 31).

The Use of Insulin to Control the Blood Glucose Level

"In patients with classic symptoms of hyperglycemia or hyperglycemia crisis, a random plasma glucose ≥ 200mg/dL (11.1 mmol/L)" (Chiang et al., 2014, p. 2036). A person with a random plasma glucose test result of less than 11.1 mmol/L (200mg/dL) is considered not to have diabetes. In a person with normal blood glucose control, their pancreas will deliver the appropriate quantity of insulin to bring their blood glucose level down to normal levels in a short period of time after the consumption of food or drink. Whereby, the pancreas of a person with T1D is incapable of producing sufficient insulin to bring their blood glucose level down in a short period of time or even in a long period of time. Therefore, the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

person with T1D has to decide whether to administer insulin to reduce their blood glucose level or to engage in some sort of physical activity to reduce their blood glucose level. A 5km walk could reduce blood glucose level by 3/4 mmol/L (54/72 mg/dL), alternatively, hoovering the house could have the same impact. Everybody is different. Therefore, blood glucose levels should be taken before and after exercise until the person knows roughly what blood glucose level to expect. "The endocrine pancreas makes insulin so that the body can use glucose (sugar) for energy. The exocrine pancreas makes enzymes that help the body digest food" (Gilbert et al., 2012, p. 37). Further research is required to investigate if insulin inhibits the release of glucose into the blood from the liver (Manninen, 2004), or if insulin enables the body to use glucose for energy or perhaps both.

Acceptable Blood Glucose Level

A person with diabetes should aim for a glucose goal of 3.9 to 6.7 mmol/L (70 to 120 mg per decilitre) before meals and a blood glucose level of no more than 10 mmol/L (180 mg per decilitre) after meals (Nathan et al., 2005). Nathan et al. suggest that a postprandial (2 hours after the meal was consumed) blood glucose level of no more than 10 mmol/L (180 mg per decilitre) while others would suggest that a postprandial blood glucose level of no more than 11mmol/L (200 mg per decilitre) is acceptable.

Unique Treatment Plan Required for Each Patient

"Type 1 diabetes seems to represent a heterogeneous disease whose pathogenic processes, genetics, and phenotypic characteristics show marked variation" (Atkinson et al., 2014, p. 70). The challenge to diabetes care teams is that each person with diabetes requires an individual treatment plan. Each person requires a specific treatment plan based on their individual lifestyle choices ((food consumption, engagement in various different exercise activities (including housework and paid work), insulin dosing regimen, stress levels, sleep patterns, and many other variables)). Unlike other medical conditions, whereby people can simply follow the dosage plan set out by their medical team, people with T1D need to plan their meals and insulin dosage numerous times a day depending on the activities with which they engage. It is important for people with diabetes to tailor the administration of insulin to suit their exercise plan to avoid episodes of hypoglycaemia (Chiang et al., 2014). Therefore, the successful treatment of people with T1D substantially depends on the person's mindfulness in medicating with insulin to reflect all that they consume, and all that they do every hour of every day. Byrne et al. (2012) suggest that people with higher educational attainment benefit more from attending the Dose Adjustment for Normal Eating (DAFNE) course when compared to people with lower educational attainment. O'Hara et al. (2016) suggest that people without skills will struggle to successfully self-manage T1D even with diabetes education and support. Leelarathna et al. (2011) speculated "that those patients with T1D who do not have access to structured education may be systematically over-insulinised" (Leelarathna et al., 2011, p. e36). Over-insulinised means the administration of too much insulin. The use of excessive insulin can lead to the patient experiencing too many hypos and severe hypos which can lead to hypo unawareness.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Pancreas

The pancreas of a person with T1D has substantially or totally lost the ability to produce insulin. The pancreas of some people after diagnosis of T1D still produce some insulin, known as residual insulin secretion (Alberti et al., 1998), this is known as the honeymoon period, but generally insufficient insulin is produced to maintain good glycaemic control. Therefore, the administration of some exogenous insulin is required, but possibly not as much insulin as would be required by a person whose pancreas no longer produces any insulin at all. Once the pancreas has lost the ability to produce insulin (after the honeymoon period), the level of glucose in the blood may rise to dangerous levels. "Diabetes mellitus is a chronic disorder of glucose metabolism with serious clinical consequences" (Forouhi & Wareham, 2010, p. 602). The most important event in the history of T1D was the discovery of insulin in 1921-22 (Atkinson et al., 2014) by Frederick G. Banting and Charles H. Best (Rotenstein et al., 2012). The discovery of insulin was not a solution to cure or prevent the onset of T1D, but the administration of exogenous insulin was an effective way to control T1D. The perennial hope that islet transplantation would alleviate the need for exogenous insulin has failed to materialise (Shapiro et al., 2000). Further research is required on islet transplantation.

Honeymoon Phase

When a patient is newly diagnosed with T1D, initially a honeymoon phase may be experienced. During the honeymoon period the body may continue to produce some insulin, therefore, less injected insulin is required to self-manage diabetes. Exercise can reduce the need for insulin as the body uses up the sugar in the blood during activity. People with T1D who engage in regular exercise may prolong the honeymoon period, maintain good glycaemic control (control of the level of glucose in the blood), and require less injected insulin for longer than newly diagnosed patients who do not engage in regular exercise. Exercise increases the body's sensitivity to insulin. Therefore, less insulin is required by people with T1D who exercise regularly, however, blood sugar levels must be monitored regularly as blood glucose levels can be very hard to predict. The extent to which the engagement in exercise can extend the honeymoon phase requires further investigation.

FACTORS THAT MAY INFLUENCE THE ONSET OF T1D

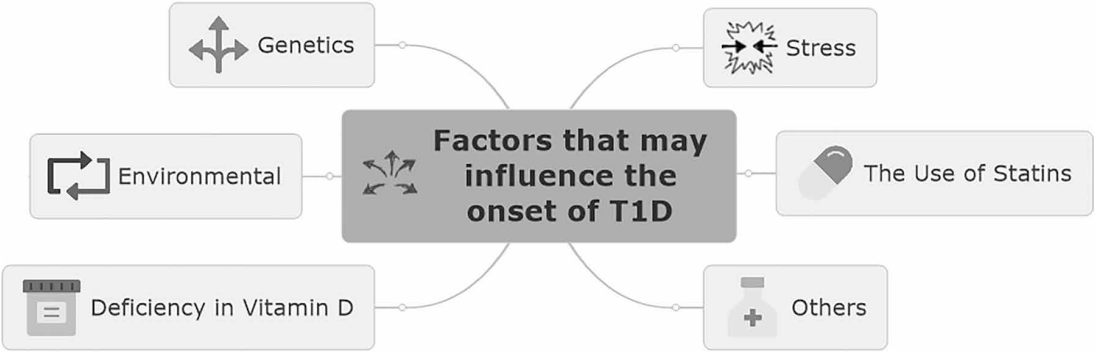
At present, there are no known ways to prevent the onset of T1D and the factors which may influence the onset are not clearly defined. Some of the factors that may influence the onset of T1D are reviewed in this section. Please refer to Figure 2.

Genetics

"Type 1 diabetes (T1D) is a common autoimmune disorder that arises from the action of multiple genetic and environmental risk factors" (Barrett et al., 2009, p. 703). Genetics can have an impact on β-cell dysfunction and death (Eizirik & Darville, 2001). "Results from linkage and association studies in T1D have long supported a model in which the major risk factor for T1D resides in the HLA region on Chromosome 6p21" (Barrett et al., 2009, p. 703). Further research is required to determine factors that trigger the onset of T1D.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Factors that may influence the onset of T1D



Environmental

"The mechanisms underlying these enigmas in geographical incidence and increased incidence rates of type 1 diabetes are unknown but have largely been attributed to environmental influences" (Atkinson et al., 2014, p. 69). Thrower and Bingley (2014) also suggest that environmental factors could play a role in the rapid increase in the diagnosis of type 1 diabetes mellitus (T1DM). Type 1 Diabetes is also known as Type 1 Diabetes Mellitus. Samuelsson and Lofman (2013) suggest that there is a possible link between the incidence of T1D and some geological factors. Further research is required to identify the specific environmental factors with possible links to the onset of T1D.

Deficiency in Vitamin D

"Vitamin D deficiency increases the risk of many common cancers, multiple sclerosis, rheumatoid arthritis, hypertension, cardiovascular heart disease, and type 1 diabetes" (Holick, 2005, p. 1024). Vitamin D is necessary in the human body to ensure that the immune system is in excellent working order. The immune system protects the body by defending it against viruses and infections. A weakened immune system can leave the body in a very vulnerable state. One such state is the diagnosis of T1D due to an immune-mediated depletion of β-cells. "Type 1 diabetes is characterised by an immune-mediated depletion of β-cells that results in lifelong dependence on exogenous insulin" (Chiang et al., 2014, p. 2034). Exogenous insulin is the insulin injected with a needle or an insulin pump. Endogenous insulin is the insulin made in the pancreas. "Type 1 diabetes (T1D) is an autoimmune disease in which immune cells, notably T Lymphocytes, target and kill the insulin secreting pancreatic beta cells" (Mahaffy & Edelstein-Keshet, 2007, p. 915). Once a large enough percentage of the beta cells have been destroyed, the pancreas stops or greatly reduces the amount of insulin produced, therefore, the glucose levels in the blood rise uncontrollably and the person is left with full blown T1D (Mahaffy & Edelstein-Keshet, 2007). From that point in time onwards, the daily administration of insulin is required for survival "to prevent the development of ketoacidosis, coma and death" (Alberti et al., 1998, p. 17). Ketoacidosis means too much sugar and ketones in the blood which may lead to DKA.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

"Numerous observational studies indicate that deficiencies in vitamins A, D, E and zinc can adversely impact immune function, particularly T-cell responses" (Kau, Ahern, Griffin, Goodman, & Gordon, 2012, p. 6). "Type 1 diabetes is a T-cell-mediated disease that is associated with loss of immunological tolerance to self-antigens" (Lindley et al., 2005, p. 92).

Although some trials have shown that the incidence of T1D in people taking Vitamin D supplements is lower than people who do not take Vitamin D supplements "Randomised controlled trials of vitamin D supplementation as a means of primary disease prevention have shown mixed results" (Thrower & Bingley, 2014, p. 682). In a study conducted by Hyppönen, Laara, Reunanen, Jarvelin, and Virtanen (2001) they concluded that their results suggested "that development of type 1 diabetes is associated with low intake of vitamin D and signs of rickets during the first year of life" (Hyppönen et al., 2001, p. 1502). The best way for the body to obtain Vitamin D is through direct contact with sunlight. During exposure to sunlight previtamin D_3 penetrates the skin, wearing sunscreen reduces the skins ability to absorb Vitamin D (Holick, 2005). The lack of vitamin D and its connection with the onset of T1D is clearly unresolved from previous studies conducted, but this interesting hypothesis is worthy of further investigation.

Stress

"As early as the beginning of the 17th century, the onset of diabetes was linked to "prolonged Sorrow" by an English Physician" (Lloyd, Smith, & Weinger, 2005, p. 121). The use of the term "Prolonged Sorrow" could possibly imply that the person may have suffered from depression. "Stress augments many hormonal and neuronal pathways, possibly over-loading pancreatic beta cells" (Butalia, Kaplan, Khokhar, & Rabi, 2016, p. 590). "Some studies have suggested that stressful experiences might affect the onset and/or the metabolic control of diabetes, but findings have often been inconclusive" (Lloyd et al., 2005, p. 121). Stressful experiences such as a difficult marriage, experiencing the grieving process alone, family breakdowns, and other circumstances may lead to the onset of T1D.

The Use of Statins

Sattar et al. (2010) found that "statin therapy was associated with a 9% increased risk for incident diabetes" (Sattar et al., p. 735). Crandall et al. (2017) also suggest that the use of statins was associated with an increased risk of developing diabetes. Casula et al. (2017) recommend rigorous monitoring of pre-diabetes patients taking statins. Further research is required to establish what causes this increased risk and how to address the risk.

EQUIPMENT, DISPOSABLES, MEDICATION, AND MONETARY COSTS

Test Strips and Lancet Devices

Blood glucose test strips are used to monitor the current level of glucose in the patients' blood. Blood glucose test strips are inserted into a meter, a lancet device is used to prick the finger to obtain a blood sample. The sample of blood is then held up to the blood glucose test strip connected to the meter which displays a number on the screen. This number indicates the level of glucose in the blood at that point

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in time only. The metre provides no indication to the patient whether or not the blood glucose level is increasing or decreasing. If tested ten minutes later, the result of a blood glucose test could be very different depending on what the person with T1D has consumed, how much insulin was administered, or the level and intensity of exercise the body has undertaken in that time. At times, a person may get an unexpected result, the blood level may be far higher or lower than expected, if this happens, the person should take a fresh sample of blood to place on a blood glucose test strip to confirm the initial level, before taking action to treat a hypo or correct for a hyper.

Ketone test strips are used to monitor the current level of ketones in the patient's blood. Ketone test strips are used with the meter in the same way as blood glucose test strips. After a blood sample has come in contact with the test strip the meter displays the level of ketones in the blood at that point in time only. After checking the blood/ketone level, the lancet (sharp) should be placed in a sharps bucket. Once a sharps bucket is full it can be dropped into a pharmacy or health clinic for safe disposal (incineration), at the same time the patient can collect an empty sharps bucket for future use. A large sharps bucket will take months to fill.

People diagnosed with T1D need to carry a diabetes medical supplies bag with them at all times. The contents of a diabetes medical supplies bag may include: insulin pens (basal and bolus), needles, meter, blood glucose test strips, lancets, diabetes diary, pen, glucose tablets, small red apple, small plastic bottle for used sharps, and whatever else is appropriate. Please refer to Figure 3.

Figure 3. Diabetes medical supplies bag

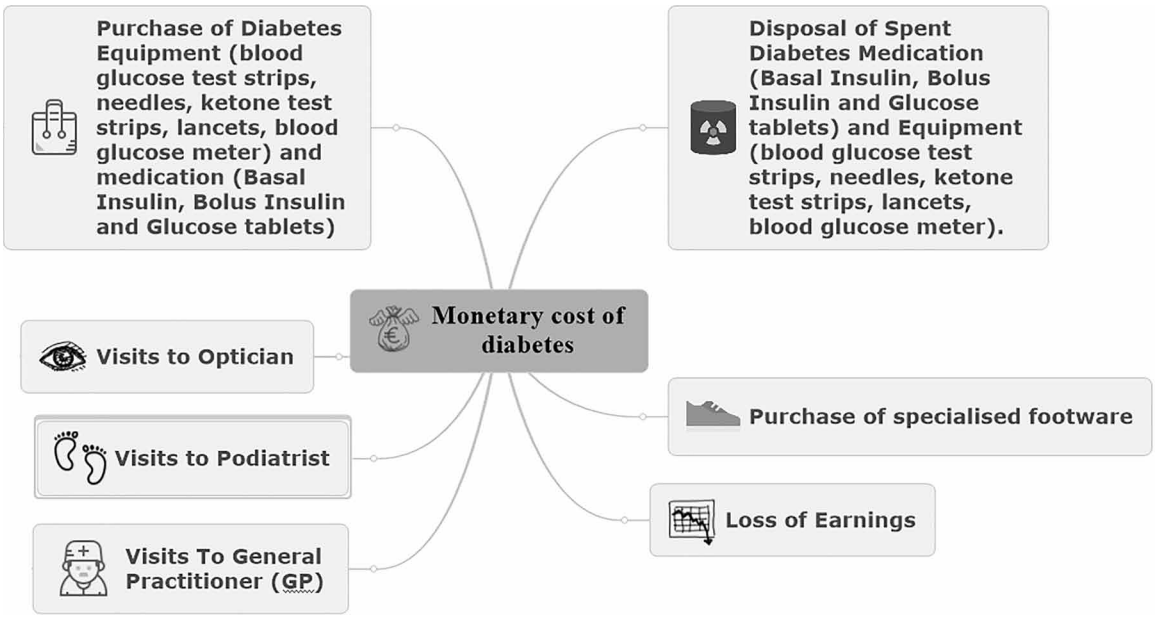
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

It is always useful to carry a small plastic bottle (for the safe collection of used sharps) and for the safe storage of sharps until the sharps can be deposited in the sharps bucket. It is the responsibility of the person to safely dispose of sharps, the doctor or diabetes care team will be able to guide the patient in this matter. Some pharmacies and health centres provide empty sharps bucket and take them back for safe disposal when full.

Monetary Costs

"Unlike their counterparts in the UK, Irish GPs are currently not reimbursed for delivering diabetes care" (O'Donnell et al., 2013, p. 3). The onset of T1D can be an expensive condition please see Figure 4. In some countries, the cost of disposable testing and treatment equipment, and medication is paid for by the health services. In other countries, the person or the family of the person has to pay for the disposable testing equipment, treatment equipment, and medication (which can be quite expensive). Regular visits to the General Practitioner (GP) are required to obtain prescriptions (every six months) to treat this life-long condition, Yearly visits to podiatrists have to be paid for in person. Expensive specialist footwear may be required to avoid complications of the feet associated with diabetes. Glasses/spectacles/contact lenses may need changing as they may no longer be suitable as the eye prescription may have changed substantially on diagnosis or may change over the years due to diabetes. Salary/wages may not be paid for the time spent in hospital or visits to medical practitioners or the diabetes care team. There may possibly be more ongoing expenses.

Figure 4. Monetary cost of diabetes



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

THE RESPONSIBLE ADMINISTRATION OF INSULIN

At present, there are no known ways to prevent the onset of T1D or to cure the condition of T1D, but it is possible to manage the condition through insulin dose adjustment and the responsible administration of insulin.

The responsible administration of insulin includes:

- Adjusting the insulin dose if necessary.
- Checking that the correct insulin pen is used: basal or bolus.
- Ensuring that the injection site is clean.
- Monitoring the injection site for changes in the skin texture or colour.
- Shoot some insulin up into the air to remove air bubbles prior to injecting insulin.
- Record the quantity of units administered.
- Confirming hypo treatments are on one's person at all times.

Basal and Bolus Insulin Injections

From the point in time of diagnoses and onwards for the rest of one's life "An immediate need for exogenous insulin replacement is also a hallmark of type 1 diabetes, for which lifetime treatment is needed" (Atkinson et al., 2014, p. 69). Insulin is generally delivered in basal (insulin is active in the body for up to 24 hours or more), and bolus (insulin is active in the body for up to four hours depending on the medication administered and the persons response to the medication) doses. A basal dose of insulin is administered to act in the background for a specific number of hours. The bolus dose of insulin is administered to cover the consumption of meals and the resulting spike in blood glucose levels over several hours. Each form of medication will have its own timeframe for dosing, and each person may respond differently to the medication, therefore, each person will have to take responsibility to monitor their blood glucose levels and discuss changes to their insulin regimen with their diabetes care team.

Basal Insulin

A basal dose of insulin is administered to act in the background all the time, while the bolus dose of insulin is administered to cover the consumption of meals and the resulting spike in blood glucose levels. The more carbohydrates in a meal; the more the blood glucose levels will rise for a number of hours after the meal has been consumed. Snorgaard, Poulsen, Andersen, and Astrup (2016) found that the more that carbohydrates were restricted in the diet the greater the reduction in blood glucose levels in a study of people with type 2 diabetes. The restriction of carbohydrates in the diets of patients with type 1 diabetes also has the effect of reducing blood glucose levels.

Bolus Insulin

A bolus dose of insulin is a rapid acting insulin which is administered at meal time (Neupane & Evans, 2014). The quantity of insulin required will depend on the carbohydrate content of the meal and the activity planned for the hours after the meal. People with T1D require training in accurately counting carbohydrates to enable them to effectively administer the correct dose of bolus (quick acting insulin).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bolus insulin should continue to lower blood glucose levels for approximately four hours after injection, depending on the brand of insulin administered. Bolus insulin is injected prior to eating, to prevent a spike in blood sugars after the consumption of food and drink. The diabetes care team should ensure that each patient understands the medication (insulin) that they are given. As with any medication: each one has individual properties and dosing schedules which should be clearly explained.

HYPERGLYCAEMIA (HYPER)

Hyperglycaemia means that the person is experiencing a high blood glucose level, which if left untreated could lead to hospitalisation, complications associated with diabetes, or death. "Current carbohydrate recommendations are based on 1) preventing ketosis, and 2) providing glucose beyond minimal needs" (Manninen, 2004, p. 9). "Ketones in the blood can lead to a condition called Ketoacidosis, also known as DKA, which just stands for diabetic ketoacidosis" (Gilbert et al., 2012, p. 21). Diabetic ketoacidosis (DKA) is life threatening, some undiagnosed, and diagnosed diabetics experience DKA to various degrees of severity. Accident and Emergency medical staff are primed to recognise the signs of DKA, and treat the patient accordingly to safely reduce the levels of glucose and ketones in the blood, and rehydrate the body.

Hyperglycaemia Episodes (Hypers)

Almost everything one consumes is converted into glucose which is released into the blood stream by the liver. Should too much glucose or sugar be consumed, this could lead to the patient going into a hyperglycaemic episode (hypers occurs when a patient's blood glucose level gets to abnormally high levels). "New hyperglycemia was defined as an admission or in-hospital fasting glucose level of 126 mg/dl (7 mmol/liter) or more or a random blood glucose level of 200mg/dl (11.1 mmol/liter) or more on 2 or more determinations" (Umpierrez et al., 2002, p. 978). Newly diagnosed hyperglycaemic patients were associated with a higher admission rate to Intensive Care Units (ICU), longer length of hospital stay, a higher in-hospital mortality rate and a lower functional outcome than previously diagnosed patients and patients with normoglycaemia (Umpierrez et al., 2002). Patients with normoglycaemia have blood sugar levels within the accepted range for the normal population (Pickup, Mattock, & Kerry, 2002). Hyperglycaemia (above normal blood sugar levels) is a dangerous medical condition which can go undetected for quite some time. "It has been thought that chronic high blood glucose levels in the body (hyperglycemia) due to diabetes could be the main cause that leads to nerve damage and thereby diabetic neuropathy" (Abdel-Motal, Abdelalim, Abou-Saleh, & Zayed, 2017, p. 172). Neuropathy can cause the feet to feel unrelenting pain (Archer, Watkins, Thomas, Sharma, & Payan, 1983). Hyperglycaemia has a key role to play in the development of diabetic neuropathies (Pop-Busui, 2010).

Treatment of Hypers

A waking blood glucose level greater than 7 mmol/L (126 mg/dL) or a two-hour postprandial blood glucose level greater than 11.1 mmol/L (200 mg/dL) could signify that the person is at risk of a hyper. A person who is experienced in managing diabetes may have the confidence to reduce their blood glucose levels by taking a correction dose of insulin, drinking plenty of water, or engaging in exercise/

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Balancing insulin administration with food consumption

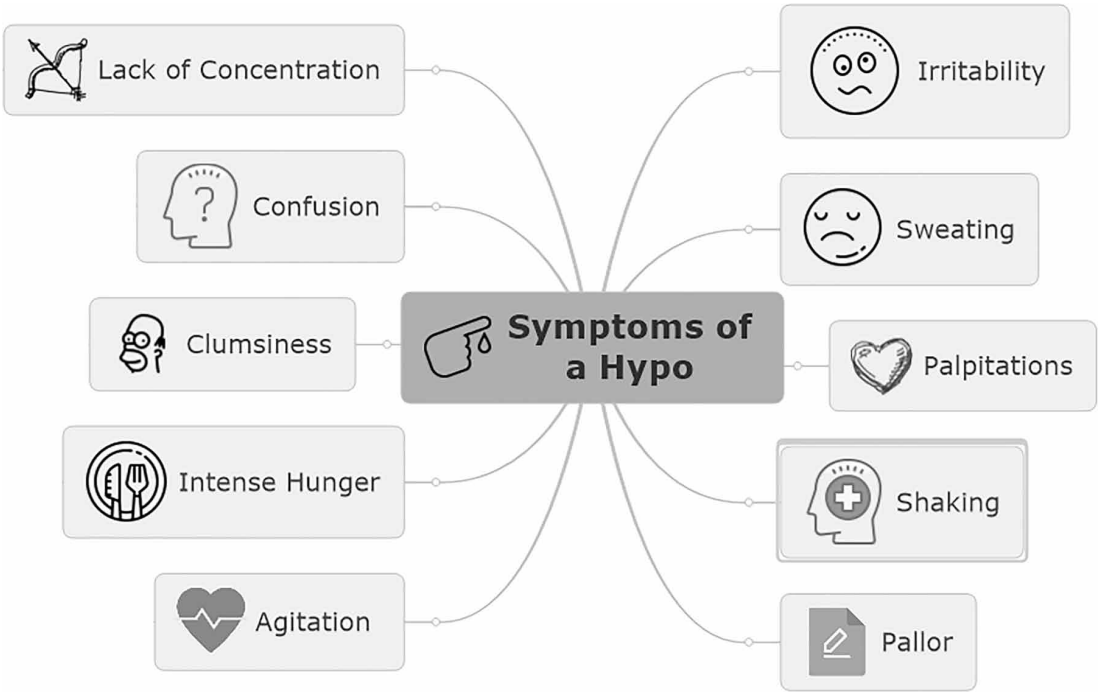
manual labour. Constant monitoring of the level of glucose in the blood is required over the next few hours, regardless of which option or combination of options are chosen to treat the high blood glucose levels. People with diabetes should consult with their diabetes care team on the treatment of hypers. A person with diabetes who is not confident to deal with a hyper themselves, should contact their Doctor/ General Practitioner, diabetes care team, or attend an Accident and Emergency (A&E) Department for treatment and guidance.

HYPOGLYCAEMIA

Calculating the correct dose of insulin to administer with a meal is a very difficult process. Please see Figure 5. Should too much bolus insulin be administered this could lead to the patient having a hypoglycaemic or severe hypoglycaemic episode over the coming hours. "The diagnosis is made by measuring the blood glucose level. If the glucose level is below 2.5mmol/l, you have hypoglycaemia" (Eldemir & Eldemir, 2011, p. 3). Hypoglycaemia is a low blood glucose level that mainly affects patients with diabetes who take insulin. Immediate consumption of glucose tablets, a piece of fruit, fruit drink, honey, or anything high in sugar is required to bring the blood glucose to a normal level and to avoid severe hypoglycaemia. Abnormally low blood glucose levels can result in hypoglycaemia which can pose a serious health threat (Byrne et al., 2012), if not effectively dealt with immediately. Should a patient with T1D on insulin not adequately treat hypoglycaemic in a relatively short time scale, this may lead to a severe hypo. "Severe hypoglycaemia was self-reported and defined as an episode of hypoglycaemia requiring the assistance of another person for treatment" (Dinneen et al., 2013, p. 31). Severe hypos should be avoided if at all possible, as a patient's ability to drive, operative machinery, or perform many other daily activities may be impaired with serious consequences should he/she suffer from a severe hypo. "Overly tight control, however, puts patients at risk of hypoglycaemic episodes (hypos) which may lead to seizures or periods of unconsciousness which may require emergency treatment" (Lawton & Rankin, 2010, p. 486). In a study conducted by Garg, Shah, Akturk, Beatson, and Snell-Bergeon (2017) one significant finding was that men had less fear of suffering from hypos than women.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Symptoms of a hypo



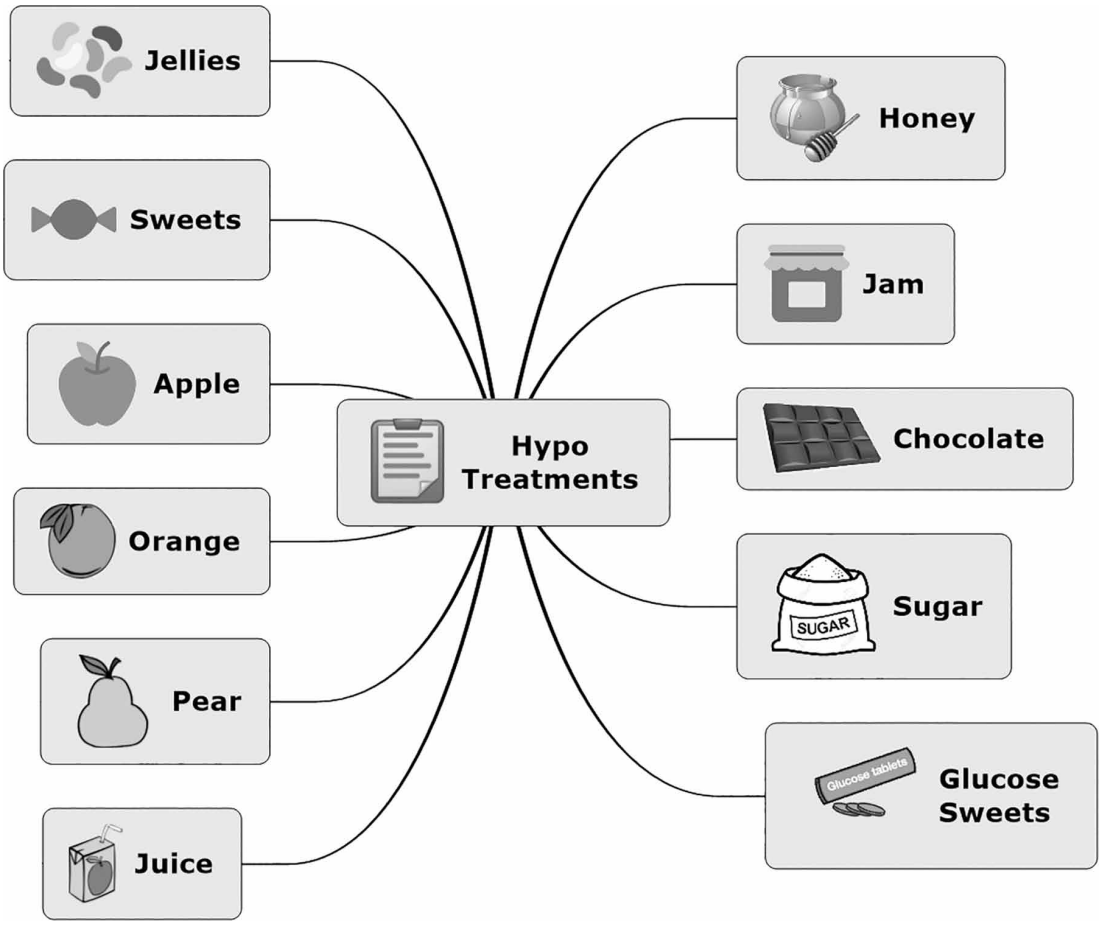
Hypoglycaemic Episodes (Hypos)

Symptoms of a hypoglycaemic episode may include some or all of the following: lack of concentration, confusion, clumsiness, intense hunger, agitation, irritability, sweating, palpitations, shaking, and pallor. Please note Figure 6. The body is basically fighting for survival as the brain feels threatened by the lack of glucose in the blood. The administration of too much insulin can inhibit the body's access to its regular glucose supply. Once the patient realises what is happening they can easily treat the hypo by taking some fast-acting glucose, for example, orange juice, a piece of fruit, or 2/3 glucose sweets. The natural instinct of a person who has just experienced a hypo is to overcompensate by the excessive consumption of sugary foods, which will invariably result in high blood glucose levels. Then the person has to decide how to deal with the high blood glucose levels. This can lead to glycaemic excursions (Deiss et al., 2006), a yo-yo effect as the blood glucose levels go up and down from one extreme to the other, as the person tries to control blood glucose levels that are uncontrollably fluctuating. People need to control the quantity of carbohydrates (carbs) consumed in the treatment of hypos which can prove to be a very challenging task.

The first experience of a hypo can take a person with diabetes on insulin by surprise, they should be alerted to this fact in advance and instructed to always carry some fast acting glucose tablets, sweets, chocolate, a piece of fruit, or fruit juice on their person. Please refer to Figure 7 for some more options for treating hypos. In addition, it is advisable to leave some glucose tablets or some other form of hypo treatment in every room of the house, including the garden shed. Gardening can be a strenuous activity that consumes the glucose in the blood which could lead to a hypo, therefore hypo treatments should always be available in the garden.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Hypo treatments



It is very easy to get locked out of the house, consequently, it is necessary to carry glucose tablets on your person everywhere. Glucose tablets should be stored in an accessible place in the car and also at the work station in the place of employment. See Figure 8 for more important locations to store hypo treatments.

RECOMMENDATIONS FOR THE SUCCESSFUL MANAGEMENT OF T1D

Some individuals with type 1 diabetes mellitus (T1DM), or T1D may be initially treated with diet alone but generally require the administration of daily doses of insulin within months of the initial diagnosis (Thrower & Bingley, 2014). The types of food and beverages consumed by the person, and the volume of activity in which the person regularly engages will influence the daily doses of insulin required to maintain good glycaemic control. The challenge with managing T1D is that every person is unique and therefore requires a unique treatment plan.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Storage locations for hypo treatments

Self-Management of Diabetes

"Despite broad organisational, intellectual, and fiscal investments, no means for preventing or curing type 1 diabetes exists, and, globally, the quality of diabetes management remains uneven" (Atkinson et al., 2014, p. 69). The self-management of T1D requires careful planning and goal planning strategies (Fredrix, McSharry, Flannery, Dinneen, & Byrne, 2018). The quality of the self-management of diabetes may depend on the educational attainment of the patient, the guidance provided by the diabetes care team, the patient's willingness to take control of the condition, diet and exercise. Not only does the quality of diabetes management remain uneven globally, but global disparities in access to insulin also must be addressed (Atkinson et al., 2014). T1D, although it predominantly affects younger members of the population, it can affect anyone at any age of their life. At present, the factors that may influence the onset of T1D are not clearly defined.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Glycaemic Control and Excursions

Some people with T1D can manage their own treatment plans, while others require substantial input from diabetes care teams to maintain good glycaemic control. People with T1D should be educated on the impact that the consumptions of fats and proteins have on glycaemic excursions. "Calculation of insulin to carbohydrate ratios (i.e. the amount of insulin needed to limit glycaemic excursions after meals) allows for increased flexibility in meal planning without sacrificing glycaemic control" (Daneman, 2006, p. 852). The Mean Amplitude of Glycaemic Excursion (MAGE) (Bando, Ebe, Muneta, Bando, & Yonei, 2017; Fritzsche, Kohnert, Heinke, Vogt, & Salzsieder, 2011; Hill et al., 2011; Syafa, Tjokroprawiro, Indra, Sargowo, & Muladi, 2014) is a method used to analyse fluctuations in the blood glucose levels of a person with diabetes. Complications associated with diabetes may be caused by glycaemic variability (Hill et al., 2011).

Glycaemic Control

Glycaemic control refers to the management of glucose levels in the blood. "Glucose control is facilitated by tracking glucose levels, care with food intake, and judging the correct does of insulin to take" (Ryan et al., 2017, p. 33). The administration of too little bolus insulin with a meal could lead to the patient's blood glucose levels reaching too high a level to be considered as good glycaemic control. The administration of too much bolus insulin with a meal could lead to the patient experiencing a hypoglycaemic episode.

HbA_{1c} Value

A HbA_{1c} value of greater than 8.5% (69 mmol/mol) would suggest that the current management of treatment by bolus and basal injections have failed the patient and other treatment methods should be considered (Neupane & Evans, 2014). The HbA_{1c} value indicates the average blood glucose level over the past two to three months, but it provides no information on the amount of time a patient may have spent in hypoglycaemic, severe hypoglycaemic, or hyperglycaemic episodes. Two patients with the same HbA_{1c} value may have experienced a wide variation in the standard deviation of blood glucose levels over the past two to three months. Fear of hypos may impede patients' ability to achieve near-normal glycaemic control. Hypos can cause death or serious injury in the immediate future. The HbA_{1c} test provides information on the probability of long-term complications. Good glycaemic control alongside avoidance of hypoglycaemic episodes would be the safest way for people with diabetes to manage their condition. "When basal-bolus routines are fastidiously applied in conjunction with the other aspects of management, many individuals with type 1 diabetes are able to maintain near-normal glycaemic control" (Daneman, 2006, p. 850).

Educational Programmes for T1D

Byrne et al. (2012) in their study of patients with T1D who participated in the DAFNE structured education programme, used several questionnaires including (i) Diabetes-Specific Quality of Life Scale (DSQOLS), (ii) Problem Areas in Diabetes Scale (PAID), and (iii) the Hospital Anxiety and Depression Scale (HADS). The conclusions to this study were that "Patients with higher baseline levels of anxiety, higher levels of diabetes related distress and higher baseline levels of HbA_{1c} are most likely to experi-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ence quality of life gain from participation in self-management programmes such as DAFNE" (Byrne et al., 2012, p. 243). Lower HbA_{1c} levels were significantly associated with adolescents who rated better Quality of Life (QOL) and lower perceived levels of family burden (Hoey et al., 2001).

Health Informatics

Fang, Pouyanfar, Yang, and Chen (2016) discuss the potential use of Big Data, Machine Learning, and Health Informatics to improve the quality of patient care. Privacy issues, ethics, validity of the data used, accessibility, security, are all relevant considerations and challenges to overcome in the use of Health Informatics. Wireless body sensor networks can be used to monitor patients' biometrics (calculation and analysis of body measurements) including blood glucose levels in healthcare (Chen, Lee, Chen, Huang, & Luo, 2009). Data mining applications can be used to analyse large volumes of data that are stored electronically in hospitals to identify trends or patterns and extract useful knowledge from the files of patients. Zaveri and Joshi (2017) suggest that data mining could be used to predict disease in advance. Patterson et al. (2009) suggest that the ability to predict potential new cases of patients with T1D over the coming years would enable European centres for diabetes treatment to plan in advance for funding requirements and health care resources to cater for the future burden of childhood diabetes. Further research is required to ascertain if the use of Big Data, Machine Learning, and Health Informatics can improve the quality of patient care.

Intensive Diabetes Therapy

"Controlled clinical trials of patients with type 1 or type 2 diabetes have not demonstrated a reduction in the occurrence of cardiovascular disease with long term intensive diabetes therapy" (Nathan et al., 2005, p. 2644). Long term intensive diabetes therapy may possibly lead to an increase in the occurrence of hypos, therefore, without substantial evidence to support the use of long-term intensive diabetes therapy, patients may be advised to keep tight control but not too intensive. "Adolescents having at least one severe hypoglycemic episode in the previous 3 months were more worried than those without" (Hoey et al., 2001, p. 1926). People should avoid having hypoglycaemic or severe hypoglycaemic episodes if at all possible; to avoid experiencing undue worry and stress, as worry can interfere with quality of life and stress can elevate blood glucose levels. Hypoglycaemic or severe hypoglycaemic episodes can be avoided by reducing the amount of insulin administered. All people with diabetes controlled by insulin should keep a diary of everything they consume in order to monitor the effects that particular foods and drinks have on their blood glucose levels. Overtime, the person will begin to know and understand their own body's response to the consumption of different foods and drinks. This knowledge should inspire confidence in the management and control of blood glucose levels and doses of insulin.

Psychosocial Outcomes

Nicolucci et al. (2013) conducted a study of diabetes-related psychosocial outcomes, 8596 adults with diabetes from 17 different countries participated in this study, a significant finding of this study was that "Approximately 40% of participants (18.6 – 64.9%) reported that their medication interfered with their ability to live a normal life" (Nicolucci et al., 2013, p. 767). Further research is required to ascertain why 40% felt that their medication interfered with their quality of life. Possibly, improved education on

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the condition would assist people in controlling their medication so that it does not interfere with their ability to live a normal life. Living in fear of hypoglycaemia, hyperglycaemia, and the possibility of diabetes related complication in the future, can be daunting.

Exercise/Physical Activity

Perhaps, if people could try to improve their understanding of diabetes, how insulin works, and how to manage exercise and general living, their quality of life may improve. Kevin Nolan (2018) a successful Gaelic Footballer with T1D hoped to inspire other people with T1D to enjoy sports. "Exercise confers remarkable benefits to those with diabetes; however, the challenge to compelling patients with diabetes to exercise are formidable" (Jenkins & Jenks, 2017, p. 968). Frequent exercise is key to controlling blood glucose levels. "Contrary to popular belief, insulin is not needed for glucose uptake and utilization in man" (Manninen, 2004, p. 7). Physical activity is one of the things that can greatly impact on glucose uptake and utilization. Therefore, patients on insulin should carefully monitor their use of insulin and blood glucose levels before, during, and for some time after physical exercise, as physical activity can increase the risk of hypoglycaemia (Seaquist et al., 2013). Patients must always reduce their insulin dose prior to engaging in exercise (Lawton & Rankin, 2010), even if it is only going for a walk or hoovering the house, as these activities can quickly consume the glucose in the blood and leave the patient at risk of experiencing a hypoglycaemic episode due to low blood glucose levels. Family members, care givers and staff should be educated on the proper surveillance and treatment of hypoglycaemia (hypos) (Seaquist et al., 2013).

Involvement in physical activities (walking, gardening, hoovering, jogging, cycling, and so forth) lowers the blood glucose levels and also has the added advantage of keeping the person away from food. Gardening may increase absorption of Vitamin D, provide valuable exercise, and reward with a beautiful display of colour as per Figure 9. Controlled consumption of food and drink alongside regular movement of the body are key to the management of blood glucose levels. Further research is required to target the formidable challenge to compel people with diabetes to participate in more exercise.

Figure 9. Picture of flowers growing in the garden

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Restricting the Consumption of Carbohydrates

Feinman et al. (2015) suggest that restricting the consumption of carbohydrates has the greatest effect on reducing blood glucose levels and can lead to the reduction or elimination of medications used in the treatment of diabetes. By cutting out or substantially reducing the consumption of foods that are high in carbohydrates (rice, pasta, white flour, potatoes, sugary drinks, cakes, biscuits and confectionery). Instead, by consuming foods that are high in nutritional content but low in carbohydrates (chicken, fish, eggs, beef, pork, full fat cheese, nuts, berries, fruit and vegetables). A person with T1D can substantially reduce the units of basal and bolus insulin administered. It would be in the interest of people living with T1D to keep a record of everything that they eat and drink for as long as it takes for them to realise the effects that different food and drink combinations have on their specific blood glucose levels. During this time, the blood glucose levels should be taken at least six times a day and recorded. Hence, the patient will be able to see exactly what effect each meal is having on their blood glucose levels. Everybody is different, therefore, it is imperative that the patient themselves or parents/carers (Silverstein et al., 2005) take responsibility for managing diabetes tailored to the patient's individual needs. "No need for an insulin head start. Low-carb, high-fat breakfasts usually allow for dosing insulin right as I start eating (or not at all)" (Brown, 2017a, p. 31). Before reducing carbohydrate consumption, the patient must reduce the units of insulin administered. "It should be recognized that the use of low-carbohydrate diets is not a recent experiment and may well approximate the diet used by much of humanity for tens of thousands of years before the rise of agriculture" (Feinman et al., 2015, p. 11). People with T1D have to be prepared to experiment with their foods, to see the effects that every morsel consumed has on blood glucose levels. The best approach is to have a low carbohydrate satisfying meal, tidy up, and walk away from the kitchen, no snacking or grazing, as this will impact on blood glucose levels. If a snack is required consume a small piece of fruit, a bag of nuts, coffee and cream, or some strawberries and cream. The consumption of full fat cream will slow down the release of natural sugars in the strawberries.

Low Carbohydrate Diet

Patients should also reduce the amount of insulin administered before commencing a low carbohydrate diet (Feinman et al., 2015). Some patients who reduced the amount of insulin administered discovered that as a result of taking control by reducing the units of insulin administered that they had also achieved much better glycaemic (blood glucose) control (Lawton & Rankin, 2010).

The Consumption of Fat and Protein Before Carbohydrates

"The temporal sequence of carbohydrate ingestion during a meal has significant impact on postprandial glucose regulation" (Shukla et al., 2017, p. 4). Shukla et al. (2017) suggest that the order in which one consumes food: fat and protein first followed by carbohydrate (carbohydrate last instead of carbohydrate first) has an impact on the absorption of the food and the resulting blood glucose level, patients that consumed carbohydrates last in a meal were found to have improved postprandial glycaemia. Postprandial glycaemia is the level of glucose in the blood two hours after consuming food.

The consumption of fat and protein with every meal provide satiety (Feinman et al., 2015), therefore, there should be no need for snacking between meals. Were a person with T1D to eat a scone with jam as a snack, the result would be a spike in blood glucose levels because there was insufficient protein and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

fat to slow down the release of glucose to the blood stream. As the spike in blood glucose levels begins to fall, the person with T1D would feel hungry again. People with diabetes should possibly avoid the consumption of scones and jam, a much healthier option would be to enjoy some fresh strawberries/blueberries/raspberries with full fat cream. Full fat cream should be consumed in moderation, as excessive consumption of full fat cream could increase total cholesterol and the amount of lipids circulating in the blood stream, which could lead to other complications.

Need to "Feed" Their Insulin

"Patients who are on insulin or insulin secretagogues are able to reduce their doses on carbohydrate-restricted diets and find they are less likely to need to "feed" their insulin" (Feinman et al., 2015, p. 6). When a person with T1D administers bolus insulin before a meal, the insulin works in the body (reducing blood glucose levels) for approximately up to four hours (everybody is different), and every medication will have a different impact on blood glucose levels. Should the quantity of insulin administered be insufficient to cover the carbohydrate content in the food consumed blood glucose levels will be higher than what is required for good glycaemic control. Should the quantity of insulin administered be in excess of what is needed to cover the carbohydrate content in the food consumed blood glucose levels will drop to a level where the patient may experience a hypoglycaemia or a severe hypoglycaemia episode. On experiencing a hypoglycaemia episode the body immediately craves food and it is difficult for the patient to restrain themselves to eating just a sufficient amount of carbohydrates to return to a safe blood glucose level. Generally, people with T1Ds will over compensate by eating too many carbohydrates after experiencing a hypoglycaemic episode. This behaviour results in the blood glucose levels escalating, feelings of guilt, and the probability that the person may then have to inject more bolus insulin to bring the blood glucose levels back down again. Frequent use of bolus insulin throughout the day can leave a person in a position whereby they are unsure as to whether their blood glucose levels are going up or down. "When blood sugars are out of range, people with diabetes are made to feel like failures" (Brown, 2017b, p. 342). Out of range blood glucose levels do not signify that the person is a failure. Possibly more information and education may be required for the person to better manage this challenging condition.

Triglycerides (TG)

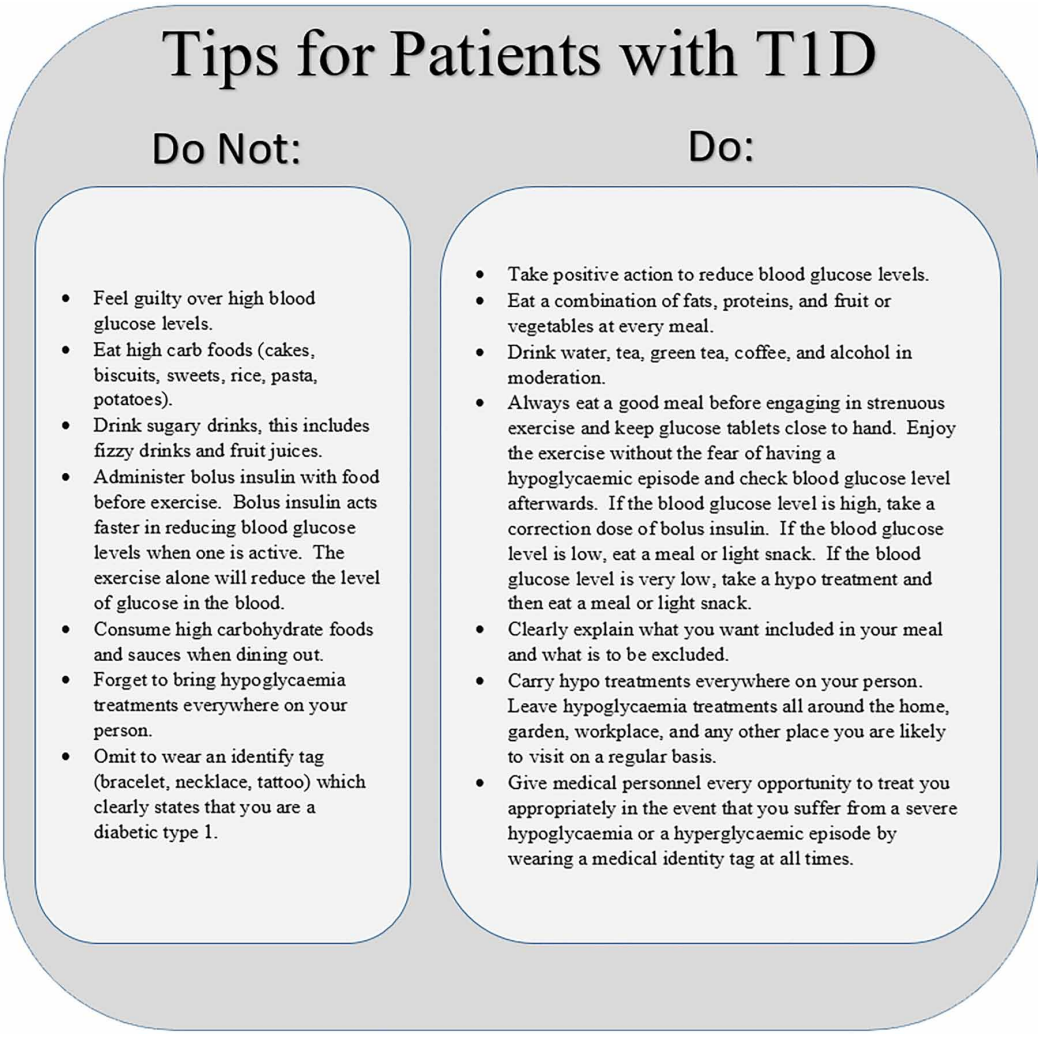
"The lower carbohydrate diets were associated with significant decreases in body weight, body mass index, TG levels, and blood pressure; additionally, they showed improvement in several other metabolic and lipid indicators" (Feinman et al., 2015, p. 7). In fact, Accurso et al. (2008) suggest that "carbohydrate restriction improves all of the features of metabolic syndrome" (Accurso et al., 2008, p. 1). Triglycerides (TG) are a type of fat/lipids found in the blood. A normal level of triglycerides found in the blood is less than 1.7mmol/L (150 milligrams per decilitre). There are many benefits to lower carbohydrate diets. There is no reason why people with T1D should not try a low carbohydrate diet for a few months while closely monitoring blood glucose levels and reducing the amount of insulin administered if they are experiencing too many hypoglycaemic episodes. People can always revert back to their normal eating patterns and insulin doses if the low carbohydrate diet does not suit them or their lifestyle. "More than a dozen additional studies, have failed to show an association between dietary lipids and risk for cardiovascular disease (CVD)" (Feinman et al., 2015, p. 7). Medina-Bravo et al. (2012) suggest that children with type 1 diabetes have higher levels of Total Cholesterol (TC) when compared to healthy subjects. "The

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

TG-to-HDL cholesterol ratio may be related to the processes involved in LDL size pathophysiology and relevant with regard to the risk of clinical vascular disease" (Boizel et al., 2000, p. 1679). High-Density Lipoproteins (HDL) are classed as the good cholesterol in the blood, as HDL transports cholesterol from all over the body back to the liver for disposal. Low-Density Lipoproteins (LDL) are classed as the bad cholesterol in the blood; as LDL can build up in the walls of blood vessels and arteries causing blockages which have the potential to cause heart attacks or strokes. The association between the consumption of dietary lipids and the risk for cardiovascular disease is very relevant to the health of people living with diabetes and the normal population, therefore, this association requires further research.

TIPS FOR PATIENTS WITH T1D

Figure 10. Tips for patients with T1D



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MEDICAL COMPLICATIONS ASSOCIATION WITH DIABETES

In the treatment of this chronic condition the person diagnosed with the condition or the parent of a child diagnosed with the condition must take immediate responsibility for administering the medication (insulin); while having limited knowledge or understanding of how diabetes and the medication works. Tight glycaemic control is required on a daily basis to avoid complications associated with diabetes in later life (Byrne et al., 2012).

Complications association with diabetes may include some or all of the following: retinopathy (Fredrix et al., 2018; Hammes, 2018; Raymakers, Gillespie, O'Hara, Griffin, & Dinneen, 2018), neuropathy (Young, Boulton, Macleod, Williams, & Sonksen, 1993), nephropathy (Andersen, Christiansen, Andersen, Kreiner, & Deckert, 1983; Chiang et al., 2014; Fowler, 2008), renal damage (O'Loughlin et al., 2010), foot ulcers (Ahmad, 2016; Fredrix et al., 2018; Noor et al., 2017), blindness, amputation (Cousart & Handley, 2016), stroke (Lawton & Rankin, 2010), cardiovascular disease (Fredrix et al., 2018; Gerstein et al., 2001; Nathan et al., 2005), myocardial infarction (heart attack), angina (chest pain or obstruction) (Nathan et al., 2005), ingrown nails (Vural et al., 2018), decrease in Bone Mineral Density (BMD) (Vestergaard, 2007), impaired ability to heal wounds (Loomans et al., 2004), immunological effects (Egan & Dinneen, 2016), proteinuria (Kiire, Horak, Lee, Klein, & Klein, 2017; Raymakers et al., 2018), anxiety and depression (Raymakers et al., 2018), and an increased risk of infection (O'Loughlin et al., 2010).

"The prevalence of each complication is generally lower with a more recent diagnosis when adjusting for duration of diabetes" (Kiire et al., 2017, p. 11). The sooner the condition is diagnosed and appropriately treated the better the prognosis (future outcomes) for the person.

Foot Ulcerations

Kelly et al. (2010) found that 30% of patients with diabetes who had been screened for foot complications tested positive for pedal neural dysfunction and 17% tested positive for vascular impairment. "Diabetic foot problems are responsible for nearly 50% of all diabetes-related hospital bed days" (Ahmad, 2016, p. 48). Therefore, it is in the patient's best interest to learn to understand how diabetes works and how best to manage the administration of insulin to control blood glucose levels; in order to avoid diabetes complications later on in life and to avoid hypoglycaemic, severe hypoglycaemic episodes, and DKA in the short term. The leading cause of hospitalisation due to diabetes is foot ulceration, foot ulcerations can adversely impact on a persons' quality of life which may lead to depression (O'Loughlin et al., 2010). "Conservative management of foot problems has dramatically reduced the risk of amputation by simple procedures, such as appropriate foot wear, cleanliness, aggressive surgical debridement, … and ulcer management" (Ahmad, 2016, p. 48). Preventative measures can be undertaken to avoid the risk of foot infection associated with diabetes and the potential for amputation (Noor et al., 2017). Comfortable appropriate footwear is necessary to avoid damage to the foot which may lead to ulceration. Frequent changing of socks during the day may be necessary to ensure that the feet are free from sweat or dampness, which could lead to fungal infections or other avoidable conditions of the feet.

People living with diabetes should monitor their feet on a daily basis by checking for dry skin, cuts, ingrown toenails, redness, swelling and/or any other abnormalities. If anything unusual is noted they should visit a podiatrist/chiropodist or their diabetes care team for a consultation and to receive advice on treating the condition. "It is concluded that acute painful diabetic neuropathy is a distinct syndrome,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

occurring in insulin or noninsulin dependent patients of any duration, and unrelated to other diabetic complications" (Archer et al., 1983, p. 491).

People with diabetes are advised to undergo a comprehensive examination of the feet every year to identify any risk factors that may lead to neuropathy or ulceration (O'Loughlin et al., 2010). Peripheral neuropathy (PN) is more common in people with type 1 diabetes who have had the condition for a longer duration (Abdel-Motal et al., 2017) and mainly affects lower limbs (Zhang et al., 2017). Diabetic peripheral neuropathy (DPN) is more likely to injure small rather than large nerve fibres (Zhang et al., 2017). "In some patients diabetic foot infection is the first presentation of diabetes and they will require diabetology input into the initial blood glucose control and ongoing diabetes management" (Sohrabi & Russell, 2017, p. 500).

Nephropathy

"There are strong associations between glucose control (as measured by haemoglobin A1c [A1C]) and the risk of developing diabetic nephropathy" (Fowler, 2008, p. 78). Nephropathy means damage or disease of the kidney which may lead to renal failure (Alberti et al., 1998). "Our study shows convincingly that diabetes nephropathy is a very severe complication of Type 1 diabetes and that survival after the onset of proteinuria is poor" (Andersen et al., 1983, p. 500). Hb is an abbreviation for haemoglobin. HbA_{1c} is an abbreviation for a blood glucose test that tells the average glycated blood glucose level over the past two to three months, there is no need for a patient to fast for a HbA1c blood test (Saudek et al., 2008). "Normal ranges for haemoglobin A1c in people without diabetes is about 4% to 5.9%" (Davis, 2016, p. 1). A HbA_{1c} of 6.5% has been recommended by an international expert committee as a cut-off point for defining overt diabetes (Dinneen, 2010). A HbA_{1c} level of 6.5% or greater has been recognised as an unmistakeable sign that the patient has diabetes.

THE CONTRAINDICATIONS OF ADMINISTERING INSULIN

"Dead in Bed" Syndrome

"Type 1 diabetes mellitus increases the risk for sudden unexplained death (SUD), generating concern that diabetes processes and/or treatments underlie these deaths. Young (<50 yrs) and otherwise healthy patients who are found dead in bed have been classified as experiencing "dead in bed" (DIB) syndrome" (Secrest, Becker, Kelsey, LaPorte, & Orchard, 2012, p. 293). Not all people with diabetes wake up on experiencing a hypoglycaemic episode during sleep, this can be dangerous as an untreated episode of hypoglycaemia can very quickly progress to an episode of severe hypoglycaemia, if a severe episode of hypoglycaemia is left untreated this could turn into a diabetic coma. Sudden Unexplained Deaths occur up to ten times more frequently in people who have been diagnosed and treated for diabetes mellitus type 1 (Secrest et al., 2012). "SUD in Type 1 DM seems to be increased 10-fold and associated with male sex, while DIB individuals have a high HbA_{1c} and insulin dose, and low BMI. Though sample size is too small for definitive conclusions, these results suggest specific sex and metabolic factors predispose to SUD and DIB" (Secrest et al., 2012, p. 293). Further research is required to establish why people with type 1 diabetes are more likely to experience SUD or DIB syndrome than people who have normal glycaemia.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Severe and Nocturnal Hypoglycaemia

There is a possibility of hypoglycaemia occurring when the level of glucose in the blood drops to ≤ 3.9 mmol/L (≤ 70 mg/dL) (Seaquist et al., 2013). Hypoglycaemia can be fatal, therefore, people who are treated with insulin should closely monitor the level of glucose in their blood on a regular basis throughout the day and always carry on their person treatments for hypoglycaemia (glucose tablets, orange juice, or a piece of fruit). Severe hypoglycaemia can be fatal (Seaquist et al., 2013). In a study conducted by Dinneen et al. (2013) 24% of respondents had experienced episodes of severe hypoglycaemia in the previous year.

People with diabetes should aim to avoid hypoglycaemia and more importantly severe hypoglycaemia episodes if at all possible during waking hours and more critically during sleep (nocturnal hypoglycaemia). "Type 1 diabetes mellitus increases the risk for sudden unexplained death (SUD), generating concern that diabetes processes and/or treatments underlie these deaths. Young (<50 yrs) and otherwise healthy patients who are found dead in bed have been classified as experiencing "dead in bed" (DIB) syndrome." (Secrest et al., 2012, p. 293).

Blood glucose concentration reduces when insulin is administered to fasting patients because it is assumed that glucose uptake into tissues is increased when insulin is administered, both Sonksen (2001) and Manninen (2004) suggest that this is just another metabolic legend and that in fact insulin lowers blood glucose levels through inhibiting hepatic glucose. The suggestion that insulin lowers blood glucose by inhibiting the livers production of glucose rather than lowering blood glucose levels by increasing glucose access into tissues, is an interesting suggestion that warrants further research.

POSITIVE IMPACT OF A DIAGNOSIS OF DIABETES

The diagnosis of T1D can initially be very confusing. Once familiar with this life-long condition, the patient then has the opportunity to take control of their condition and manage their blood glucose levels to suit their lifestyle and also to maintain good health for as long as possible.

Opportunity to Lead a Very Healthy Life

The consumption of too many sugary foods can have detrimental effects on everybody whether they suffer with diabetes or not. Through regular finger pricking, people with diabetes can see the impact that all foods and drinks have on their blood glucose levels. Therefore, they have the choice to consume foods and beverages which will not spike their blood sugars, and hence live a much healthier life by cutting out the consumption of foods and beverages that have high sugar content and little nutritional value.

Weight Loss

The sudden onset of T1D can lead to weigh loss. Some people struggle with their weight for years and never lose weight for long. Many attend various different classes to lose weight, lose a stone or two over a period of time, and then put it all back on (and frequently more weight). So, when a person who is newly diagnosed with T1D has lost weight as a result of their condition, they then have the opportunity to keep the weight off, by managing the condition, exercising, and eating good nutritious food.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Frequent Access to Excellent Medical Professionals

Once diagnosed as a person with T1D, depending on the country of domicile, the patient will be assigned to a diabetes care team. As with everything in life, some diabetes care teams will be better funded and resourced than others. But still the opportunity is there for the patient to tap into the knowledge base of the medical professionals to ensure that they effectively manage their condition while maintaining an excellent quality of life. The diabetes care team perhaps see each patient twice a year. Therefore, blood testing would be conducted twice a year. The results of the blood tests, (analysed by the diabetes care team), give a very clear picture of the overall general health of the patient. Hence, patients with diabetes have access to excellent medical professionals on a regular basis to guide them to a healthier lifestyle. People who are not diagnosed with diabetes, do not get regular blood tests and access to medical teams to guide them on how to improve their overall health, unless, they pay for these regular health check-ups.

Constant Reminder to Eat Well and Be Active

Even though the excessive consumption of carbohydrates and sugars may not spike the blood glucose levels of people without diabetes this may still lead to other unhealthy medical conditions, for example, obesity. "As a guide for the patient with diabetes, the prescription of many agencies to "eat to the meter" seems like a good one" (Feinman et al., 2015, p. 11). The blood glucose reading on the meter can be a wake-up call to the patient; if the reading is higher than the patient would like perhaps they have not engaged in sufficient exercise or work to expend the glucose created from their last meal, so perhaps they should go for a walk, do some housework or gardening prior to the consumption of more food or drink (with the exception of water). Many members of the world's population currently eat more calories per day than their bodies require.

Some people continually graze (eat and drink unconsciously), the effective management of diabetes depends on one's ability to control every morsel of food or drink consumed. Or, at least to record everything consumed, measure the impact on the blood glucose levels and manage the blood glucose levels to bring them to an acceptable level by engaging in exercise or through the administration of an appropriate dose of insulin.

Self-Management of Condition Unlike Other Conditions

In a study conducted by Nicolucci et al. (2013) a notable finding was that "27.7% of respondents reported a positive impact of diabetes on at least one aspect of their life" (Nicolucci et al., 2013, p. 772). Of all the diseases or conditions to get, T1D is possibly one of the easiest to self-manage, providing the patient is sufficiently informed to manage the condition, willing to monitor every morsel consumed, check blood glucose levels regularly, and take responsibility for altering the doses of (basal and bolus) insulin to suit their lifestyle. The management of T1D in babies, children, adolescents, disabled people, and the elderly is totally different to self-management of the condition; as the person responsible for the administration of insulin has no feedback from their own body to warn in advance of hypos. The management of T1D in babies, children, adolescents, disabled people, and the elderly, is deserving of further research.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

A number of interesting questions have arisen during this review of T1D which deserve further investigation:

- To investigate if insulin inhibits the release of glucose into the blood from the liver (Manninen, 2004), or if insulin enables the body to use glucose for energy or perhaps both.
- Islet transplantation.
- The extent to which engagement in exercise can extend the honeymoon phase.
- The determination of factors that trigger the onset of T1D.
- The identification of environmental factors with possible links to the onset of T1D.
- The lack of vitamin D and its connection to the onset of T1D.
- The use of statins linked to an increased risk of developing diabetes.
- To ascertain if the use of Big Data, Machine Learning, and Health Informatics can improve the quality of patient care.
- To investigate why people diagnosed with T1D felt that their medication interfered with their quality of life.
- To target the formidable challenge to compel people with diabetes to participate in more exercise.
- The association between the consumption of dietary lipids and the risk for cardiovascular disease.
- To establish why people with T1D are more likely to experience SUD or DIB syndrome than people who have normal glycaemia.
- The management of T1D in babies, children, adolescents, disabled people, and the elderly.

CONCLUSION

The diagnosis of T1D, a debilitating autoimmune disease, is now considered to be a global epidemic. The self-management of T1D is a challenging ongoing process requiring daily focus to maintain good glycaemic control. For the successful treatment of this life-long medical condition the patient or carer requires an understanding of this condition. Self-management of diet, engagement in physical activity, and the daily dosage of insulin administered, are necessary to avoid complications associated with diabetes in the future. One of the main challenges to people with T1D is to avoid hypers and hypos. Hypers occur when the blood glucose levels get too high and hypos occur when the blood glucose levels get too low. Both conditions are potentially dangerous to the patient. Therefore, the successful management of diabetes is for the patient to keep their blood glucose levels under tight control at all times.

A HbA_{1c} result between 4% and 5.6%, a fasting blood glucose level less than 7 mmol/L (126 mg/dl), and an Oral Glucose Tolerance Test (OGTT) below 11.1 mmol/L (200mg/dL) are the levels expected for normoglycaemia (members of the population not affected by diabetes). A person diagnosed with T1D should aim for blood glucose levels near normoglycaemia while avoiding hypos. The quality of the self-management of the condition may depend on the educational attainment of the patient, the expertise of the diabetes care team, and by the patient attending and engaging with diabetes educational programmes. An awareness of simple facts, such as, the consumption of fats and proteins before carbohydrates can delay the release of glucose into the blood and result in a lower postprandial (2 hours after food consumption)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

blood glucose level. People with T1D are 10 times more likely to suffer from Sudden Unexplained Deaths (SUD) and Dead-in-Bed (DIB) syndrome. It is possible for a person with T1D to control their condition to suit their lifestyle, and therefore possibly avoid complications associated with diabetes in later life.

CONFLICTS OF INTEREST

There are no potential conflicts of interest relevant to this article.

FUNDING

This research did not receive any grant from funding agencies.

REFERENCES

Abdel-Motal, U., Abdelalim, E., Abou-Saleh, H., & Zayed, H. (2017). Neuropathy of type 1 diabetes in the Arab world: A systematic review and meta-analysis. *Diabetes Research and Clinical Practice*, *127*, 172–180. doi:10.1016/j.diabres.2017.03.001 PMID:28384559

Accurso, A., Bernstein, R., Dohlqvist, A., Draznin, B., Feinman, R., Fine, E., ... Vernon, M. (2008). Dietary carbohydrate restriction in type 2 diabetes mellitus and metabolic syndrome: Time for a critical appraisal. *Nutrition & Metabolism*, *5*(9), 1–8. PMID:18397522

Ahmad, J. (2016). The diabetic foot. *Diabetes & Metabolic Syndrome*, *10*(1), 48–60. doi:10.1016/j.dsx.2015.04.002 PMID:26072202

Ahmed, N., Choe, Y., Mustad, V., Chakraborty, S., Goates, S., Luo, M., & Mechanick, J. (2018). Impact of malnutrition on survival and healthcare utilization in Medicare beneficiaries with diabetes: A retrospective cohort analysis. *BMJ Open Diabetes Research & Care*, *6*(1e000471), 19. doi:10.1136/bmjdrc-2017-000471 PMID:29449950

Alberti, K., Aschner, P., Assal, J., Bennett, P., Groop, L., Jervell, J., . . . Zimmet, P. (1998). *Definition, diagnosis and classification of diabetes mellitus and its complications - Report of WHO Consultation*.

Andersen, A., Christiansen, S., Andersen, J., Kreiner, S., & Deckert, T. (1983). Diabetic nephropathy in type 1 (insulin-dependent) diabetes: an epidemiological study. *Diabetologia*, *25*, 496–501.

Ang, L., Jaiswal, M., Callaghan, B., Raffel, D., Brown, M., & Pop-Busui, R. (2017). Sudomotor dysfunction as a measure of small fiber neuropathy in type 1 diabetes. *Autonomic Neuroscience: Basic and Clinical*, *205*, 87–92. doi:10.1016/j.autneu.2017.03.001 PMID:28325598

Archer, A., Watkins, P., Thomas, P., Sharma, A., & Payan, J. (1983). The natural history of acute painful neuropathy in diabetes mellitus. *Journal of Neurology, Neurosurgery, and Psychiatry*, *46*(6), 491–499. doi:10.1136/jnnp.46.6.491 PMID:6875582

Atkinson, M., & Eisenbarth, G. (2001). Type 1 diabetes: new perspectives on disease pathogenesis and treatment. *The Lancet, 358*(9277), 221-229.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Atkinson, M., Eisenbarth, G., & Michels, A. (2014). Type 1 diabetes. *Lancet*, *383*(9911), 69–82. doi:10.1016/S0140-6736(13)60591-7 PMID:23890997

Bando, H., Ebe, K., Muneta, T., Bando, M., & Yonei, Y. (2017). Proposal for Insulinogenic Index (IGI)-Carbo70 as experimental evaluation for diabetes. *Journal of Clinical and Experimental Endocrinology*, *1*(1), 1–5.

Barrett, J., Clayton, D., Concannon, P., Akolkar, B., Cooper, J., Erlich, H., ... Rich, S. (2009). Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nature Genetics*, *41*(6), 703–707. doi:10.1038/ng.381 PMID:19430480

Boizel, R., Laporte, F., Benhamou, P., Foulon, T., Lardy, B., & Halimi, S. (2000). Ratio of triglycerides to HDL cholesterol is an indicator of LDL particle size in patients with type 2 diabetes and normal HDL cholesterol level. *Diabetes Care*, *23*(11), 1679–1685. doi:10.2337/diacare.23.11.1679 PMID:11092292

Brown, A. (2017a). *Bright spots & landmines: the diabetes guide I wish someone had handed me*. The diaTribe Foundation.

Brown, A. (2017b). Fried chicken, an airport ride, and why diabetes is so hard. *Clinical Diabetes, American Diabetes Association*, *35*(5), 340–344. doi:10.2337/cd17-0095 PMID:29263579

Butalia, S., Kaplan, G., Khokhar, B., & Rabi, D. (2016). Environmental risk factors and type 1 diabetes: Past, present and future. *Canadian Journal of Diabetes*, *40*(6), 586–593. doi:10.1016/j.jcjd.2016.05.002 PMID:27545597

Byrne, M., Newell, J., Coffey, N., O'Hara, M., Cooke, D., & Dinneen, S. (2012). Predictors of quality of life gains among people with type 1 diabetes participating in the Dose Adjustment for Normal Eating (DAFNE) structured education programme. *Diabetes Research and Clinical Practice*, *98*(2), 243–248. doi:10.1016/j.diabres.2012.09.017 PMID:23018180

Byrne, M., O'Connell, A., Egan, A., Dinneen, S., Hynes, L., O'Hara, M., ... Coyne, I. (2017). A core outcomes set for clinical trials of interventions for young adults with type 1 diabetes: An international, multi-perspective Delphi consensus study. *Bio Med Central*, *18*(602), 1–8. PMID:29258565

Casula, M., Mozzanica, F., Scotti, L., Tragni, E., Pirillo, A., Corrao, G., & Catapano, A. L. (2017). Statin use and risk of new-onset diabetes: A meta-analysis of observational studies. *Nutrition, Metabolism, and Cardiovascular Diseases*, *27*(5), 396–406. doi:10.1016/j.numecd.2017.03.001 PMID:28416099

Chen, S., Lee, H., Chen, C., Huang, H., & Luo, C. (2009). Wireless body sensor network with adaptive low-power design for biometrics and healthcare applications. *IEEE Systems Journal*, *3*(4), 398–409. doi:10.1109/JSYST.2009.2032440

Chiang, J., Kirkman, M., Laffel, L., & Peters, A. L. (2014, July). Type 1 diabetes through the life span: A position statement of the American Diabetes Association. *Diabetes Care*, *37*(7), 2034–2054. doi:10.2337/dc14-1140 PMID:24935775

Cousart, T., & Handley, M. (2016). Implementing diabetic foot care in the primary care setting. *The Journal for Nurse Practitioners, 13*(3), e129-e132.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Crandall, J., Mather, K., Rajpathak, S., Goldberg, R., Watson, K., Foo, S., ... Temprosa, M. (2017). Statin use and risk of developing diabetes: Results from the Diabetes Prevention Program. *BMJ Open Diabetes Research & Care*, *5*(1e000438), 1–8. doi:10.1136/bmjdrc-2017-000438 PMID:29081977

Daneman, D. (2006). Type 1 diabetes. *The Lancet, 367*(9513), 847-858.

Davis, C. (2016). Hamoglobin A1c Test (HbA1c, A1c, Hb1c). Retrieved from https://www.medicinenet.com/hemoglobin_a1c_test/article.htm#hemoglobin_a1c_definition_and_facts

Deiss, D., Bolinder, J., Riveline, J., Battelino, T., Bosi, E., Tubiana-Rufi, N., ... Philip, M. (2006). Improved glycemic control in poorly controlled patients with Type 1 Diabetes using real-time continuous glucose monitoring. *Diabetes Care*, *29*(12), 2730–2732. doi:10.2337/dc06-1134 PMID:17130215

Dinneen, S. (2010). What is diabetes? *Diabetes: Basic facts*. *Medicine*, *38*(11), 11. doi:10.1016/j.mpmed.2010.08.004

Dinneen, S., O'Hara, M., Byrne, M., Smith, D., Courtney, C., McGurk, C., ... O'Shea, D. (2013). Group follow-up compared to individual clinic visits after structured education for type 1 diabetes: A cluster randomised controlled trial. *Diabetes Research and Clinical Practice*, *100*(1), 29–38. doi:10.1016/j.diabres.2013.01.017 PMID:23398978

Egan, A., & Dinneen, S. (2014). What is diabetes? *Diabetes: Basic Facts*. *Medicine*, *42*(12), 12. doi:10.1016/j.mpmed.2014.09.005

Egan, A., & Dinneen, S. (2016). In-hospital metabolic regulation in patients with a diabetic foot ulcer: Is it worthwhile? *Diabetes/Metabolism Research and Reviews*, *32*(Suppl. 1), 297–302. doi:10.1002/dmrr.2741 PMID:26453180

Eizirik, D., & Darville, M. (2001). B-cell apoptosis and defense mechanisms: lessons from type 1 diabetes. *Diabetes, 50*(Supplement 1), S64-S69.

Eldemir, G., & Eldemir, E. (2011). Hypoglycaemia (low blood glucose) in non-diabetic people: what can cause hypoglycaemia episodes in non-diabetic patients? Retrieved from http://www.netdoctor.co.uk/conditions/a4460/hypoglycaemia-low-blood-glucose-in-non-diabetic-people/

Fang, R., Pouyanfar, S., Yang, Y., & Chen, S. (2016). Computational Health Informatics in the Big Data Age: a survey. *ACM Computing Surveys, 49*(1), 12:11-12:36.

Feinman, R., Pogozelski, W., Astrup, A., Bernstein, R., Fine, E., Westman, E., ... Worm, N. (2015). Dietary carbohydrate restriction as the first approach in diabetes management: Critical review and evidence base. *Nutrition (Burbank, Los Angeles County, Calif.)*, *31*(1), 1–13. doi:10.1016/j.nut.2014.06.011 PMID:25287761

Forouhi, N., & Wareham, N. (2010). Epidemiology of diabetes. *Diabetes: Basic facts*. *Medicine*, *38*(11), 11. doi:10.1016/j.mpmed.2010.08.007

Fowler, M. (2008). Microvascular and macrovascular complications of diabetes. *Clinical Diabetes*, *26*(2), 77–82. doi:10.2337/diaclin.26.2.77

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fredrix, M., McSharry, J., Flannery, C., Dinneen, S., & Byrne, M. (2018). Goal-setting in diabetes self-management: A systematic review and meta-analysis examining content and effectiveness of goal-setting interventions. *Psychology & Health*, *33*(8), 955–977. doi:10.1080/08870446.2018.1432760 PMID:29498547

Fritzsche, G., Kohnert, K., Heinke, P., Vogt, L., & Salzsieder, E. (2011). The use of computer program to calculate the mean amplitude of glycemic excursions. *Diabetes Technology & Therapeutics*, *13*(3), 319–325. doi:10.1089/dia.2010.0108 PMID:21291337

Garg, S., Shah, V., Akturk, H., Beatson, C., & Snell-Bergeon, J. (2017). Role of mobile technology to improve diabetes care in adults with Type 1 Diabetes: The remote-T1D study iBGStar in Type 1 Diabetes management. *Diabetes Therapy, Springer*, *8*(4), 811–819. doi:10.100713300-017-0272-5 PMID:28555339

Gerstein, H., Mann, J., Yi, Q., Zinman, B., Dinneen, S., Hoogwerf, B., ... Yusuf, S. (2001). Albuminuria and risk of cardiovascular events, death, and heart failure in diabetic and nondiabetic individuals. *Journal of the American Medical Association*, *286*(4), 421–426. doi:10.1001/jama.286.4.421 PMID:11466120

Gilbert, K., Millard, R., & Flynn, G. (2012). *Type 1 diabetes starter kit: a guide for adults with type 1 diabetes*: Diabetes Ireland and Type 1 Diabetes Network Australia.

Griesdale, D., deSouza, R., vanDam, R., Heyland, D., Cook, D., Malhotra, A., ... Talmor, D. (2008). Intensive insulin therapy and mortality among critically ill patients: A meta-analysis including NICE-SUGAR study data. *Canadian Medical Association Journal*, *180*(8), 821–827. doi:10.1503/cmaj.090206 PMID:19318387

Hammes, H. (2018). Diabetic retinopathy: Hyperglycaemia, oxidative stress and beyond. *Diabetologia*, *61*(1), 29–38. doi:10.100700125-017-4435-8 PMID:28942458

Hex, N., Bartlett, D., Wright, D., Taylor, M., & Varley, D. (2012). Estimating the current and future costs of type 1 and type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine*, *29*(7), 855–862. doi:10.1111/j.1464-5491.2012.03698.x PMID:22537247

Hill, N., Oliver, N., Choudhary, P., Levy, J., Hindmarsh, P., & Matthews, D. (2011). Normal reference range for mean tissue glucose and glycemic variability derived from continuous glucose monitoring for subjects without diabetes in different ethnic groups. *Diabetes Technology & Therapeutics*, *13*(9), 921–928. doi:10.1089/dia.2010.0247 PMID:21714681

Hoey, H., Aanstoot, H., Charelli, F., Daneman, D., Danne, T., Dorchy, H., ... Aman, J. (2001). Good metabolic control is associated with better quality of life in 2,101 adolescents with type 1 diabetes. *Diabetes Care*, *24*(11), 1923–1928. doi:10.2337/diacare.24.11.1923 PMID:11679458

Holick, M. (2005). Vitamin D: Important for prevention of osteoporosis, cardiovascular heart disease, type 1 diabetes, autoimmune diseases, and some cancers. *Southern Medical Journal*, *98*(10), 1024–1027. doi:10.1097/01.SMJ.0000140865.32054.DB PMID:16295817

Hurley, L., O'Donnell, M., O'Hara, M., Carey, M., Willaing, I., Daly, H., & Dinneen, S. (2017). Is diabetes self-management education still the Cinderella of diabetes care? *Patient Education and Counseling*, *100*(10), 1957–1960. doi:10.1016/j.pec.2017.05.026 PMID:28545846

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hyppönen, E., Laara, E., Reunanen, A., Jarvelin, M., & Virtanen, S. (2001). Intake of vitamin D and risk of type 1 diabetes: a birth-cohort study. *The Lancet, 358*(9292), 1500-1503.

Jackson, C., Albanese-O'Neill, A., Butler, K., Chiang, J., Deeb, L., Hathaway, K., ... Siminerio, L. (2015). Diabetes care in the school setting: A position statement of the American Diabetes Association. *Diabetes Care*, *38*(10), 1958–1963. doi:10.2337/dc15-1418 PMID:26404925

Jenkins, D., & Jenks, A. (2017). Exercise and diabetes: A narrative review. *The Journal of Foot and Ankle Surgery*, *56*(5), 968–974. doi:10.1053/j.jfas.2017.06.019 PMID:28842107

Kau, A., Ahern, P., Griffin, N., Goodman, A., & Gordon, J. (2012). Human nutrition, the gut microbiome, and immune system: Envisioning the future. *Nature*, *474*(7351), 327–336. doi:10.1038/nature10213 PMID:21677749

Kelly, S., Dolan, C., Hurley, L., Kelly, L., Garrow, A., O'Shea, E., ... Dinneen, S. (2010). Prevalence of diabetic foot complications in the west of Ireland: A pilot study. *The Diabetic Foot Journal*, *13*(2), 82–89.

Kiire, C., Horak, K., Lee, K., Klein, B., & Klein, R. (2017). The period effect in the prevalence of proliferative diabetic retinopathy, gross proteinuria, and peripheral neuropathy in type 1 diabetes: A longitudinal cohort study. *PLoS One*, *12*(3), 1–14. doi:10.1371/journal.pone.0174979 PMID:28362881

Lawton, J., & Rankin, D. (2010). How do structured education programmes work? An ethnographic investigation of the dose adjustment for normal eating (DAFNE) programme for type 1 diabetes patients in the UK. *Social Science & Medicine*, *71*(3), 486–493. doi:10.1016/j.socscimed.2010.04.030 PMID:20621746

Leelarathna, L., Ward, C., Davenport, K., Donald, S., Housden, A., Finucane, F., & Evans, M. (2011). Reduced insulin requirements during participation in the DAFNE (dose adjustment for normal eating) structured education programme. *Diabetes Research and Clinical Practice*, *92*(2), e34–e36. doi:10.1016/j.diabres.2011.01.001 PMID:21269721

Lindley, S., Dayan, C., Bishop, A., Roep, B., Peakman, M., & Tree, T. (2005). Defective suppressor function in CD4+ CD25+ t-cells from patients with type 1 diabetes. *Diabetes*, *54*(1), 92–99. doi:10.2337/diabetes.54.1.92 PMID:15616015

Lloyd, C., Smith, J., & Weinger, K. (2005). Stress and diabetes: A review of the links. *Diabetes Spectrum*, *18*(2), 121–127. doi:10.2337/diaspect.18.2.121

Loomans, C., deKoning, E., Staal, F., Rookmaaker, M., Verseyden, C., deBoer, H., ... Zonneveld, A. (2004). Endothelial progenitor cell dysfunction: A novel concept in the pathogenesis of vascular complications of type 1 diabetes. *Diabetes*, *53*(1), 195–199. doi:10.2337/diabetes.53.1.195 PMID:14693715

Mahaffy, J., & Edelstein-Keshet, L. (2007). Modeling cyclic waves of circulating T cells in autoimmune diabetes. *SIAM Journal on Applied Mathematics is the property of Society for Industrial and Applied Mathematics, 67*(4), 915-937.

Malone, J., Pavan, P., Morrison, A., & Cuthbertson, D. (2001). Prevalence and significance of retinopathy in subjects with type 1 diabetes of less than 5 years' duration screened for the diabetes control and complications trial. *Diabetes Care*, *24*(3), 522–526. doi:10.2337/diacare.24.3.522 PMID:11289479

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Manninen, A. (2004). Metabolic effects of the very-low-carbohydrate diets: Misunderstood "villians" of human metabolism. *Journal of the International Society of Sports Nutrition*, *1*(2), 7–11. doi:10.1186/1550-2783-1-2-7 PMID:18500949

Medina-Bravo, P., Medina-Urrutia, A., Juarez-Rojas, J., Cardoso-Saldana, G., Jorge-Galarza, E., Posadas-Sanchez, R., ... Posadas-Romers, C. (2012). Glycemic control and high-density lipoprotein characteristics in adolescents with type 1 diabetes. *Pediatric Diabetes*. PMID:23057424

Nathan, D., Cleary, P., Backlund, J., Saul, M., Genuth, M., Lachin, J., ... Zinman, B. (2005). Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *The New England Journal of Medicine*, *353*(25), 2643–2653. doi:10.1056/NEJMoa052187 PMID:16371630

Neupane, S., & Evans, M. (2014). Modern strategies for management of glycaemia in type 1 diabetes. Medicine, *42*(12), 707-710. Retrieved from https://www.sciencedirect.com/science/article/pii/S1357303914002709

Nicolucci, A., Burns, K., Holt, R., Comaschi, M., Hermanns, N., Ishii, H., ... Peyrot, M. (2013). Diabetes attitudes, wishes and needs second study (DAWN2TM): Cross-national benchmarking of diabetes-related psychosocial outcomes for people with diabetes. *Diabetic Medicine*, *30*(7), 767–777. doi:10.1111/dme.12245 PMID:23711019

Nolan, K. (2018). Kevin Nolan - Winning against the odds. Retrieved from https://www.diabetes.ie/living-with-diabetes/living-type-1/kevin-nolan/

Noor, S., Khan, R., & Ahmad, J. (2017). Understanding diabetic foot infection and its management. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *11*(2), 149–156. doi:10.1016/j.dsx.2016.06.023 PMID:27377687

O'Donnell, M., deSuin, A., O'Mullane, M., Smith, D., Bradley, C., Finucane, F., & Dinneen, S. (2013). Difference in the structure of outpatient diabetes care between endocrinologist-led and general physician-led services. *BMC Health Services Research*, *13*(493), 1–9. PMID:24274036

O'Hara, M., Hynes, L., O'Donnell, M., Keighron, C., Allen, G., Caulfield, A., ... Dinneen, S. (2017). Strength in numbers: An international consensus conference to develop a novel approach to care delivery for young adults with type 1 diabetes, the D1 Now Study. *Research Involvement and Engagement*, *3*(25), 1–13. PMID:29214056

O'Hara, M., Hynes, L., O'Donnell, M., Nery, N., Byrne, M., Heller, S., & Dinneen, S. (2016). Systematic review of meta-analysis: A systematic review of interventions to improve outcomes for young adults with type 1 diabetes. *Diabetic Medicine*, 34(6), 753–769. PMID:27761951

O'Loughlin, A., McIntosh, C., Dinneen, S., & O'Brien, T. (2010). Basic concepts to novel therapies: A review of the diabetic foot. *The International Journal of Lower Extremity Wounds*, *9*(2), 90–102. doi:10.1177/1534734610371600 PMID:20483808

Patterson, C., Dahlquist, G., Gyurus, E., Green, A., Soltesz, G., & Group, E. S. (2009). Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: A multicentre prospective registration study. *Lancet*, *373*(9680), 2027–2033.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Pickup, J., Mattock, M., & Kerry, S. (2002). Glycaemic control with continuous subcutaneous insulin infusion compared with intensive insulin injections in patients with type 1 diabetes: Meta-analysis of randomised controlled trials. *British Medical Journal*, *324*, 1–6. PMID:11909787

Pop-Busui, R. (2010). Cardiac autonomic neuropathy in diabetes: A clinical perspective. *Diabetes Care*, *33*(2), 434–441. doi:10.2337/dc09-1294 PMID:20103559

Raymakers, A., Gillespie, P., O'Hara, M., Griffin, M., & Dinneen, S. (2018). Factors influencing health-related quality of life in patients with type 1 diabetes. *Health and Quality of Life Outcomes*, *16*(27), 1–5. PMID:29394942

Rotenstein, L., Kozak, B., Shivers, J., Yarchoan, M., Close, J., & Close, K. (2012). The ideal diabetes therapy: What will it look like? How close are we? *Clinical Diabetes*, *30*(2), 44–53. doi:10.2337/diaclin.30.2.44

Russell, S., El-Khatib, F., Sinha, M., Magyar, K., McKeon, K., Goergen, L., ... Damiano, E. (2014). Outpatient glycemic control with a bionic pancreas in type 1 diabetes. *The New England Journal of Medicine*, 371(4), 313–325. PMID:24931572

Ryan, E., Holland, J., Stroulia, E., Bazelli, B., Babwik, S., Li, H., ... Greiner, R. (2017). Improved A1c levels in type 1 diabetes with smartphone app use. *Canadian Journal of Diabetes*, *41*(1), 33–40. doi:10.1016/j.jcjd.2016.06.001 PMID:27570203

Samuelsson, U., & Lofman, O. (2013). Geochemical correlates to type 1 diabetes incidence in Southeast Sweden: An environmental impact? *National Environmental Health Association*, *76*(6), 146–154. PMID:24645426

Sattar, N., Preiss, D., Murray, H., Welsh, P., Buckley, B., de Craen, A., ... Ford, I. (2010). Statins and risk of incident diabetes: A collaborative meta-analysis of randomised statin trials. *Lancet*, *375*(9716), 735–742. doi:10.1016/S0140-6736(09)61965-6 PMID:20167359

Saudek, C., Herman, W., Sacks, D., Bergenstal, R., Edelman, D., & Davidson, M. (2008). A new look at screening and diagnosing diabetes mellitus. *The Journal of Clinical Endocrinology and Metabolism*, *93*(7), 2447–2453. doi:10.1210/jc.2007-2174 PMID:18460560

Seaquist, E., Anderson, J., Childs, B., Cryer, P., Dagogo-Jack, S., Fish, L., ... Vigersky, R. (2013, May). Hypoglycemia and Diabetes: A report of a workgroup of the American Diabetes Association and The Endocrine Society. *Diabetes Care*, *36*(5), 1384–1395. doi:10.2337/dc12-2480 PMID:23589542

Secrest, A., Becker, D., Kelsey, S., LaPorte, R., & Orchard, T. (2012). Characterising sudden death and dead-in-bed syndrome in type 1 diabetes: Analysis from 2 childhood-onset type 1 diabetes registries. *Diabetic Medicine*, *28*(3), 293–300. PMID:21309837

Shapiro, A., Lakey, J., Ryan, E., Korbutt, G., Toth, E., Warnock, G., ... Rajotte, R. (2000). Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *The New England Journal of Medicine*, *343*(4), 230–238. doi:10.1056/NEJM200007273430401 PMID:10911004

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shukla, A., Andono, J., Touhamy, S., Casper, A., Lliescu, R., Mauer, E., ... Aronne, L. (2017). Carbohydrate-last meal pattern lowers postprandial glucose and insulin excursions in type 2 diabetes. *BMJ Open Diabetes Research & Care*, *5*(1e000440), 1–5. doi:10.1136/bmjdrc-2017-000440 PMID:28989726

Silverstein, J., Klingensmith, G., Copeland, K., Plotnick, L., Kaufman, F., Laffel, L., ... Clark, N. (2005). Care of children and adolescents with type 1 diabetes - A statement if the American Diabetes Association. *Diabetes Care*, *28*(1), 186–212. doi:10.2337/diacare.28.1.186 PMID:15616254

Snorgaard, O., Poulsen, G., Andersen, H., & Astrup, A. (2016). Systemic review and meta-analysis of dietary carbohydrate restriction in patients with type 2 diabetes. *BMJ Open Diabetes Research & Care*, *5*(e000354), 1–10.

Sohrabi, S., & Russell, D. (2017). Diabetic foot and foot debridement technique. *Surgery, 35*(9), 500-504.

Sonksen, P. (2001). Hormones and Sport: Insulin, growth hormone and sport. *The Journal of Endocrinology*, *170*(1), 13–25. doi:10.1677/joe.0.1700013 PMID:11431133

Spallone, V., Ziegler, D., Freeman, R., Bernardi, L., Frontoni, S., Pop-Busui, R., ... Valensi, P. (2011). Cardiovascular autonomic neuropathy in diabetes: Clinical impact, assessment, diagnosis, and management. *Diabetes/Metabolism Research and Reviews*, *27*(7), 639–653. doi:10.1002/dmrr.1239 PMID:21695768

Syafa, L., Tjokroprawiro, A., Indra, J., & Sargowo, D., & Muladi. (2014). Expert system for blood glucose fluctuations measurement based on MAGE (Mean Amplitude of Glycemic Excursion) and HbA1c on diabetic using K-NN (Nearest Neighbor). *Journal of Basic and Applied Scientific Research*, *4*(12), 135–141.

Thrower, S., & Bingley, P. (2014). What is type 1 diabetes? *Diabetes: Basic Facts. Medicine*, *42*(12), 12. doi:10.1016/j.mpmed.2014.09.003

Umpierrez, G., Isaacs, S., Bazargan, N., You, X., Thaler, L., & Kitabchi, A. (2002). Hyperglycemia: An independent marker of in-hospital mortality in patients with undiagnosed diabetes. *The Journal of Clinical Endocrinology and Metabolism*, *87*(3), 978–982. doi:10.1210/jcem.87.3.8341 PMID:11889147

Vestergaard, P. (2007). Discrepancies in bone mineral density and fracture risk in patients with type 1 and type 2 diabetes - a meta-analysis. *International Osteoporosis Foundation and National Osteoporosis Foundation*, *18*(4), 427–444. doi:10.100700198-006-0253-4 PMID:17068657

Voltarelli, J., Couri, C., Stracieri, A., Oliveira, M., Moraes, D., Pieroni, F., ... Burt, R. (2007). Autologous nonmyeloablative hematopoietic stem call transplantation in newly diagnosed type 1 diabetes mellitus. *Journal of the American Medical Association*, *297*(14), 1568–1576. doi:10.1001/jama.297.14.1568 PMID:17426276

Vural, S., Bostanci, S., Koçyigit, P., Çaliskan, D., Baskal, N., & Aydin, N. (2018). Risk factors and frequency of ingrown nails in adult diabetic patients. *The Journal of Foot and Ankle Surgery*, *57*(2), 289–295. doi:10.1053/j.jfas.2017.10.006 PMID:29329712

Wen, L., Ley, R., Volchkov, P., Stranges, P., Avanesyan, L., Stonebraker, A., ... Chervonsky, A. (2008). Innate immunity and intestinal microbiota in the development of type 1 diabetes. *Nature*, *455*(7216), 1109–1113. doi:10.1038/nature07336 PMID:18806780

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Young, M., Boulton, A., Macleod, A., Williams, D., & Sonksen, P. (1993). A multicentre study of the prevalence of diabetic peripheral neuropathy in the United Kingdom hospital clinic population. *Diabetologia*, *36*(2), 150–154. doi:10.1007/BF00400697 PMID:8458529

Zaveri, S., & Joshi, N. (2017). A comparative study of data analysis techniques in the domain of medicative care for disease prediction. *International Journal of Advanced Research in Computer Science*, *8*(3), 564–566.

Zhang, X., Fang, C., Li, X., Cao, Y., Zhang, Q., Zhang, H., ... Liu, C. (2017). Clinical characteristics and risk factors of diabetic peripheral neuropathy of type 1 diabetes mellitus patients. *Diabetes Research and Clinical Practice*, *129*, 97–104. doi:10.1016/j.diabres.2017.04.016 PMID:28521198

This research was previously published in the International Journal of User-Driven Healthcare (IJUDH), 8(2); edited by Rakesh Biswas and Ashok Kumar Biswas; pages 14-45; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 3
For Better Healthcare Mining Health Data

Güney Gürsel
Command post of Gendarme Logistics, School ol of Technical and Auxiliary Forces, Turkey

ABSTRACT

Data mining has great contributions to the healthcare such as support for effective treatment, healthcare management, customer relation management, fraud and abuse detection and decision making. The common data mining methods used in healthcare are Artificial Neural Network, Decision trees, Genetic Algorithms, Nearest neighbor method, Logistic regression, Fuzzy logic, Fuzzy based Neural Networks, Bayesian Networks and Support Vector Machines. The most used task is classification. Because of the complexity and toughness of medical domain, data mining is not an easy task to accomplish. In addition, privacy and security of patient data is a big issue to deal with because of the sensitivity of healthcare data. There exist additional serious challenges. This chapter is a descriptive study aimed to provide an acquaintance to data mining and its usage and applications in healthcare domain. The use of Data mining in healthcare informatics and challenges will be examined.

INTRODUCTION

From primary care institutions to big healthcare centers, every healthcare organization uses an information system. These healthcare information systems (HCIS) store, process and retrieve healthcare data. Healthcare data is very valuable in today's world. By the help of rapidly developing healthcare informatics, there are efforts to use the valuable data stored electronically in HCIS databases to improve healthcare. Healthcare staff expects more than e-recording the data from HCISs. Besides using healthcare data for care giving, healthcare centers and academic centers use these data for education and research. Research in medical area is not limited to healthcare development such as developing new healing techniques and drugs. There are healthcare informatics fields such as structured data entry, constructing longitudinal patient data, image processing, etc. Any research area dealing with huge and valuable data such as medical domain requires creative techniques supported with computers and computer systems to utilize it. Data mining techniques are good examples for these required creative techniques.

DOI: 10.4018/978-1-7998-1204-3.ch003

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Life expectancy has risen. We live longer when we compared to past. This improvement is tremendous for human being. But it has a cost. We spend more to healthcare when compared to past. The healthcare expenses are becoming unmanageable. Payers put pressure on healthcare organizations to decrease the cost. Then is it possible to decrease the cost while increasing the quality of healthcare? The answer is "it depends". It depends on the extent that we utilize the data and information we have.

It is stated that every healthcare institution uses an information system. During the daily healthcare service, every piece of patient data is recorded by means of HCISs. In short, we have huge amounts of patient data and it is growing every day. These huge amounts of data are called as "Big

Data". These "Big Data" should be used to get information and knowledge. The most fundamental challenge about Big Data is to explore the large volumes of data and extract useful information or knowledge for future use (Wu, Zhu, Wu, & Ding, 2014).

Data mining is one of the techniques to utilize this valuable healthcare data. There are many different definitions available in the literature. Drawing from the literature, we can define data mining as the process of analyzing data and discovering knowledge, patterns, associations, rules, anomalies, sequences that are non-trivial, implicit, previously unknown and potentially useful from databases.

Healthcare institutions are geographically distributed to ease healthcare access. Data about a patient can be scattered into many HCIS databases. There should be a technique to handle these scattered data to analyze and mine. Distributed data mining deals with this scattered data problem and will be covered through the chapter.

Patient health data is very sensitive. Researches about patient health data should comply with some ethical and legal issues to protect patient privacy as well as data mining studies. Privacy is term used for notion of the confidentiality and access restrictions of patients' protected health information (PHI) which contains sensitive and personal information (Sun, Zhu, Zhang, & Fang, 2012). Patient privacy refers to any medical data (medical condition, test result, payment information, etc.) in any form (paper, electronic, etc.) that belongs to a patient, the meaning is about what is protected and who will be permitted to use this information (Upstate Medical University, 2011). Protecting patient privacy in data mining will be a part and it is examined in the chapter.

The complexity and toughness of the medical environment is common information in the literature. The complexity and toughness of the medical domain is a challenge to every technology and application as well as data mining. In this chapter, the challenges specific to data mining will also be examined.

This chapter is a descriptive study that examines the concepts, issues related and techniques used in data mining in healthcare big data. The purpose of the study is to give an idea about the data mining tasks and techniques used in healthcare.

The chapter is organized as follows: In the background section, the universal basics, not related to healthcare only, of data mining will be covered. The main part is in the heading of "MEDICAL DATA MINING" comprising the contribution of the data mining to the healthcare, data mining techniques used in healthcare, temporal data mining in healthcare, distributed data mining in healthcare, challenges, privacy and security issues, data mining for healthcare management. In the end, are the future research directions and conclusion parts interpreting the chapter.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

Data Mining Basics

In this section, the universal basics, not related to healthcare only, of data mining will be examined. There will be a brief summary of data mining to give idea and get acquaintance. Not to digress from the main subject, details are left to reader; the references in the "Additional Reading" could help.

Data mining is the miracle of technology era. In the electronically stored big data, no matter what area is, companies are in a race to get useful knowledge to gain advantage in the ruthless competition of our age. To give an idea about what can data mining extract from the big data, we can give some examples as (Bramer, 2013):

- Analyzing satellite imagery,
- Analysis of organic compounds,
- Fraud detection,
- Electric load prediction,
- Financial forecasting,
- Medical diagnosis,
- Real estate valuation,
- Targeted marketing,
- Toxic hazard analysis,
- Weather forecasting.

In data mining, the critical part is the quality of data. Before data mining, we have to be sure that data has below features (Kantardzic, 2011):

- **Accuracy:** The data to be mined should be consistent, complete and true.
- **Having Right Data Type:** We have to ensure that the data is stored in the correct form. Numerical value should not be available in character form or vice versa.
- **Integrity:** There shouldn't be any conflicts in the data.
- **Consistency:** The form and the content should not change after integration of large data sets.
- **Being Not Redundant:** Redundancy is an issue that should be avoided or kept to minimum not only in data mining but also in database design. It is the replication of the same data in different tables of database. Redundant data may lead to problems in storing, accessing and analyzing the data.
- **Timeliness:** The time component of the data should be well organized and bias related to data time should be avoided.
- **Being Understandable:** The data should not have misleading form. The user and analyzer should clearly understand what the data stand for.
- **Complete:** Missing values are serious problems for data mining. It should also be kept to minimum.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The features given above are very critical. In real world, the data in the databases are so messy to deal with which is called as "dirty". They are mostly incomplete, erroneous (noisy) and inconsistent. In order to make the data have desired features given above, data preprocessing is done for data mining. Major tasks in data preprocessing can be listed as (Jiawei & Kamber, 2001):

- **Data Cleaning:** Filling missing values, smoothing noisy data, identifying or removing outliers and resolving inconsistencies.
- **Data Integration:** Integration of multiple databases, data cubes or files.
- **Data Transformation:** Normalization and aggregation.
- **Data Reduction:** Obtains reduced representation in volume but produces the same or similar analytical results.
- **Data Discretization:** Part of data reduction but with particular importance especially for numerical data.

Data mining tasks and techniques are confused and can be misused interchangeably. To eliminate the confusion, it is better to mention about what are data mining tasks and what data mining techniques are.

Data mining tasks are divided into two categories: Descriptive tasks and Predictive tasks. Some sources in literature name these categories as Supervised and Unsupervised. Descriptive tasks present the main characteristics of the data set, a summary of the data points making it possible to study important aspects (Jain & Srivastava, 2013). Descriptive tasks are undirected, find patterns in the data set and leave the interpretation of the patterns to the researcher in which the data speaks for itself (Jain & Srivastava, 2013).

Predictive tasks provide the researcher with the facility to predict an unknown (often future) value of a specific variable (Jain & Srivastava, 2013).

Namely, data mining tasks are:

- Classification,
- Clustering,
- Rule association,
- Description,
- Visualization,
- Outlier detection,
- Time series analysis.

In Classification, we try to find group memberships for the known and predefined labels (classes). Classification is a predictive data mining task.

In Clustering, there is no predefined and known class. Clustering tasks group the data into clusters according to their similarities.

Rule association is another predictive task that is used for searching relationships within data. Antecedent and consequent relationships, meaning which occurs before and after, which situations co-exist, etc. are the possible search criteria.

Description tasks are descriptive and they help us characterize the data. By the help of description tasks, data can be generalized, summarized and characterizations are determined.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Visualization tasks help us understand the results of data mining. Because some data mining techniques are hard to understand for the target managers, by means of visualization, result of data mining becomes accessible and comparable.

Outlier detection is easy to understand from its name. These tasks help us detect the data that are significantly different, we call them as outlier. Especially for fraud and abuse detection, outlier detection tasks are used.

Time series analysis tries to find patterns and rules depending on time.

As to data mining techniques, we can define them as the methods used to perform data mining tasks given above. The common techniques for data mining can be listed as:

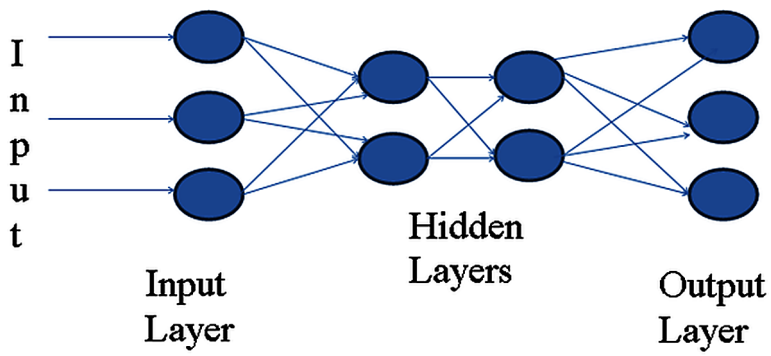
- Artificial Neural Network (ANN),
- Decision Tree (DT),
- Genetic Algorithms (GAs),
- Regression and Correlation methods (such as Logistic Regression),
- Bayesian Networks,
- Support Vector Machines (SVMs),
- Similarity Measures (such as Nearest Neighbor Method),
- Fuzzy Logic,
- Hybrid techniques (such as Neuro-Fuzzy).

Artificial Neural Network (ANN) is the model that tries to emulate human nerve system into computer applications. ANN is an adaptive system that changes its parameters during the training phase, like human nerve system, model learns from the training set. In Figure 1, an ANN design example is given. Number of hidden layers depends on the case and the problem.

Decision Tree (DT) is a structure where each node denotes a test on an attribute value, each branch represents an outcome of the test and tree leaves represent classes or class distribution, it is a predictive model used for classification commonly (Jain & Srivastava, 2013). In Figure 2 an example of DT is given.

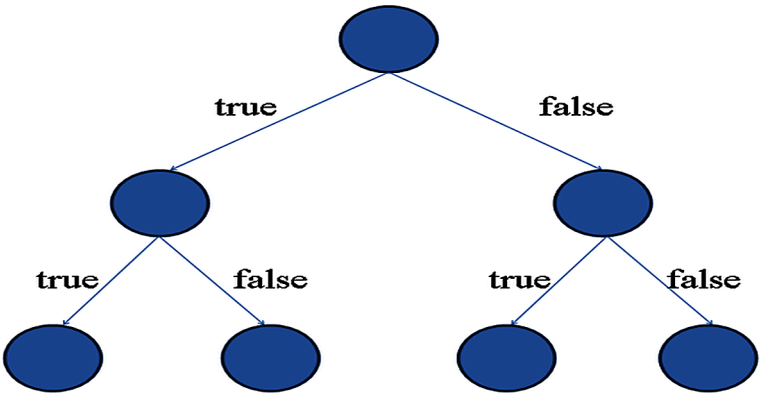
Genetic Algorithms approach is the effort to emulate natural evaluation into computer applications. The idea behind GAs is building a better solution by combining the good parts of other solutions just like nature does by combining the DNA of living beings (Jain & Srivastava, 2013).

Figure 1. ANN example



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. DT example



Regression and Correlation methods (such as Logistic Regression) evaluates the strength of a relationship where regression is a predictive method of future values by fitting the current available values into a curve, and correlation is a descriptive method used to examine the degree of similarity (Jiawei & Kamber, 2001).

Bayesian Network is a statistical multivariate model for a set of variables which is defined in terms of two components, a qualitative component, a directed acyclic graph (DAG), and a quantitative component, a conditional distribution of each variable in the graph(Jensen & Nielsen, 2007).

Support Vector Machines are manly used for classification. It is a type of supervised machine learning method. The method is presented with a set of labeled data instances and the SVM training algorithm aims to find a hyper plane separating the given dataset into a discrete predefined number of classes consistent with the training set (Mountrakis, Im, & Ogole, 2011).

Similarity Measures are used for measuring the similarities in big data. Nearest Neighbor Method, Manhattan distance, and Euclidean distance are the most known forms of similarity measures.

Fuzzy logic is the multivalued logic attempting to emulate human reasoning in the computer applications. In fuzzy logic, instead of using hard borders such as 0-1, white-black, soft computing techniques are employed.

Hybrid techniques combine fuzzy logic methods with other methodologies such as neural network, genetic algorithms, Bayesian reasoning, machine learning, etc.

MEDICAL DATA MINING

Contribution of the Data Mining to the Healthcare

Healthcare service can have huge costs. Food consumed, bed occupied and related services for lodging, surgery preparation laboratory tests, radiology tests, surgery, post-surgery care and post-surgery control laboratory tests... expenditures for the inpatients increase the healthcare cost in a considerable manner. Payers of the healthcare put strong pressure on the healthcare institutions to decrease the costs. By the way, the toughest thing is the quality of the service must increase in spite of the decreased cost not to decrease the attractiveness of the healthcare institution.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

USA spent 2 trillion dollars in 2008 for healthcare (which stands for 16.2% of gross domestic product) (AMA, 2010). Table 1 gives the top-ranked countries in healthcare expenditure (Speights, 2013).

As seen in Table 1, the expenditures are huge. These expenditures are not compatible with the life expectancy, USA is the first to spend but near last to have life expectancy whereas Switzerland is almost last to spend but first to have life expectancy.

American Medical Association's (AMA, 2010) proposal to reduce costs can be summarized as follows:

- Reduce the burden of preventable disease.
  - Reduction of the risk factors for disease;
  - Prevention of the onset of chronic illness;
  - Improvement of patient compliance with medications.
- Reduce nonclinical health system costs that do not contribute to patient care
- Make health care delivery more efficient.
  - Reduction of unnecessary use of services;
  - Increase in use of services with positive return on investment;
  - Increase in availability of information on the effectiveness of different treatments;
  - Improvement in management of chronic illness;
  - Reduction of medical errors.
- Promote value-based decision-making at all levels.

Achievement of all proposed solutions of AMA is directly dependent on the use of information and conversion of the information to knowledge.

Above, we mention that there is a pressure to decrease the costs, but at the same time the healthcare institutions have to increase the quality. Is this possible? The answer is "it depends". It depends on the extent that the managements of the healthcare institutions utilize the big data they have. There are some utilization techniques; one of them is data mining.

Although the discussion about what is data, what is information and what is knowledge took place long ago and these terms are clear in today's world, it will be useful to mention briefly in this chapter to avoid confusion. Data is the raw form and it has no meaning by itself. Wikipedia (2015) explains

Table 1. Countries with the highest healthcare expenditure in OECD

Rank	Country	Percentage in Gross National Product (%)	$ per Person Spent	Life Expectancy (Years)
1	USA	17.9	8,680	78.7
2	Netherlands	12	5,056	80.8
3	Germany	11.6	4,338	80,5
4	France	11.6	3,974	81,3
5	Switzerland	11.4	5,270	82.6
6	Canada	11.4	4,445	80.8
7	Denmark	11.1	4,464	79.3

Source: Speights, 2013.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the relation between these three concepts in simple and understandable way. Difference between these three terms lies in the level of abstraction. Data is the first level, the raw and meaningless form. With interpreting, the data becomes information. Utilization of information brings the knowledge.

Returning back to our subject, it is said that almost all healthcare institutions use an information system. These systems collect and generate data. These data are not only about patient but also about medical devices, medical resources, their operating times, in short, every item related to healthcare. The healthcare data at hand are growing bigger and bigger as the time passes. The anonymous saying (it is everywhere but no sign about who is the owner) "Drowning in data, starving for information/knowledge" summarizes the situation. We need data mining to extract information and knowledge from these big data in healthcare, use this information and knowledge for decreasing the costs and increasing the quality.

Data mining help healthcare get (Hays, 2012):

- Question based answers,
- Anomaly based discovery,
- New Knowledge discovery,
- Informed decisions,
- Probability measures,
- Predictive modeling,
- Decision support,
- Improved health,
- Personalized medicine.

In the above items, it is clear that the profits of medical data mining are compatible with the proposed solutions of AMA given before. They are very tough to accomplish. All needs long time effort, discrete planning and powerful techniques.

In addition to Hay's list, Canlas (2009) added Evidence based medicine, Reduction of medical errors, Policy making in public health, Early diagnosis, Preventive medicine, Non-invasive diagnosis, Detection of adverse drug events (Canlas, 2009).

Other additional items to this list made by Desikan, Hsu, and Srivastava (2011) are:

- Fraud and abuse detection,
- Customer relationship management decisions of Healthcare organizations,
- Identification of effective treatments and best practices by physicians,
- More affordable healthcare services received by patients.

Association rules are the important aid of the data mining to the healthcare. It gives characteristics of certain information as a consequence or as an antecedent (Milovic & Milovic, 2012). This is very helpful for health professionals who are searching for the relations between diseases and lifestyles or demographics or between survival rates and treatment (Milovic & Milovic, 2012). The associations are used to help strengthen the claims regarding whether to engage or eliminate the rules in the knowledge model (Milovic & Milovic, 2012).

One of the most value adding contributions of data mining to the healthcare is in reimbursement field. Medicare, the national social insurance program administered by US government, declared that it

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

wouldn't reimburse the healthcare organization for medical errors in 2009. This would lead serious cuts in the healthcare organizations. Reduction of medical errors, by means of data mining, not only gives way to more qualified service, but also helps elimination of the deserved fund cuts.

In addition to elimination of fund cuts, data mining also enables organizations to have accurate claims to the reimbursement companies. By the use of the information and knowledge discovered, underestimation or over billing risks will definitely be lower.

Data Mining Techniques Used in Healthcare Informatics

The data mining techniques/methods and data mining tasks may be a point of confusion in the literature as it is mentioned before. To avoid misunderstanding, before starting the section, it is better to clarify that in this chapter, classification, clustering, rule association, description and visualization are called as data mining tasks. The techniques/methods used accomplishing these tasks such as Artificial Neural Networks, Decision Tree, Logistic Regression etc. will be called as data mining techniques/methods.

In healthcare, HCISs record huge amount of data every day. But quality of the data decreases because of the complex nature of healthcare. It is previously mentioned that data should be clean to apply successful data mining applications. There are rigorous data mining tools to handle noisy data. But anyway the healthcare institutions should take measures to capture and store complete, consistent and accurate data for better data mining results.

Healthcare is an endless area for the data mining. So it is hard to tell any technique, not used in it. In this section of the chapter, the most common techniques will be examined and explained briefly.

The most common task used in medical data mining is the classification (Chen, Fuller, Friedman, & Hersh, 2005).

One of the application areas is the detection of tumors in radiology. Radiologists have difficulties in detection of tumors. The neural networks with back-propagation and association rule mining is commonly used for tumor classification in mammograms (Mishra, Padhy, & Panigrahi, 2013). The data mining is effectively used in the diagnosis of lung abnormality that may be cancerous or benign (Mishra, Padhy, & Panigrahi, 2013).

For the detection of high risk breast cancer groups, it is proved possible to find statistically significant associations with breast cancer by deriving a decision tree and selecting the best leaf (Sundar, Latha, & Chandra, 2012).

Statistical neural networks are effectively used for breast cancer diagnosis by applying several neural network structures (Sarvestani, Safavi, Parandeh, & Salehi, 2010).

In the Sundar, Latha, and Chandra's study (2012), results showed that SVM has better results for diabetes and heart disease whereas decision tree is better technique for liver disorders and breast cancer.

In the literature review of Kolçe and Frasheri (2012) for the data mining techniques used in the diagnosis and prognosis of diseases, DTs, ANNs and Bayesian algorithms appeared to be the most well-performing algorithms for diagnosis. For prognosis ANNs, Bayesian algorithms, DTs and Fuzzy algorithms came out to be the most well-performing algorithms respectively. In the same study, literature showed that, in diagnosis of cancer diseases ANNs, in heart diseases Bayesian algorithms and in other diseases DTs, in prognosis of cancer diseases again ANNs, in heart diseases ANNs and Bayesian algorithms are the most well performing techniques.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Temporal Data Mining in Healthcare Informatics

Time is a critical dimension in healthcare knowledge extraction because hidden patterns do not exist throughout the whole time period of healthcare monitorization but only in some time intervals with recurrent and periodic nature (Henriques, Pina, & Antunes, 2013).

Temporal data mining (TDM) can be defined as the attempt to search for interesting correlations and patterns in temporal databases (Bruno & Garza, 2012). TDM tries to discover qualitative and quantitative patterns in temporal databases or in datasets of discrete-valued time series (Lin, Orgun, & Williams, 2002). It has two popular threads (Lin, Orgun, & Williams, 2002):

- In the purpose of finding fully or partially similar patterns in temporal databases(similarity)
- In the purpose of finding fully or partially periodic patterns in temporal databases(periodicity)

Temporal Abstraction is the approach of dealing with time-related data in medical research in which methods give the opportunity to make abstract definition of temporal data by extracting the relevant features (Takabayashi, Ho, Yokoi, Nguyen, Kawasaki, Le, Suziki, & Yokosuka, 2007). It is the transformation of time-stamped data into interval based representation. Temporal Abstraction has two phases:

1. First one is extracting time-stamped data,
2. Second is extracting temporal specific relationships between these extracted data (Takabayashi, Ho, Yokoi, Nguyen, Kawasaki, Le, Suziki, & Yokosuka, 2007).

Batal, Valizadegan, Cooper & Hauskrecht (2012), has divided the clinical variable as medical variables (that uses the abstraction as **on** medication and **of** medication) and lab variables having two types of abstractions, trend abstraction (decreasing, steady, increasing) and value abstraction (e.g. very low, low, high, etc.).

To describe the relations between pairs of state intervals, Allen's temporal logic (1984) is used in TDM not in medicine but in general. Table 2 shows Allen's 7 relations between the time intervals A and B.

One of the most attractive applications of TDM is the extraction of temporal rules from data (Laboratory for Biomedical Informatics, 2015). In temporal rules, the consequent is related to the antecedent of the rule with a kind of temporal relationship, unlike association rules; in addition, a temporal rule suggests a cause-effect association between the antecedent and the consequent of the rule (Laboratory for Biomedical Informatics, 2015).

For example, by the help of TDM, a temporal rule extracted may be, "in many diabetic patients, the presence of hyperglycemia overlaps with the absence of glycosuria" (Patel, Hsu, & Lee, 2008). Another example can be "Low systolic blood pressure AND Low diastolic blood pressure frequently occur before high heart rate frequency" (Concaro, Sacchi,Cerra, Fratino, & Bellazzi, 2011).

Distributed Data Mining in Healthcare Informatics

Healthcare organizations are (and should be) distributed to facilitate easy access of public. In most countries, there is a chain of healthcare that is; the patient visits the primary care health institutions first, if it can't solve the problem, then secondary and third step health institutions are visited respectively. This type of healthcare organization brings some problems for healthcare informatics such as standard-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Allen's temporal logic

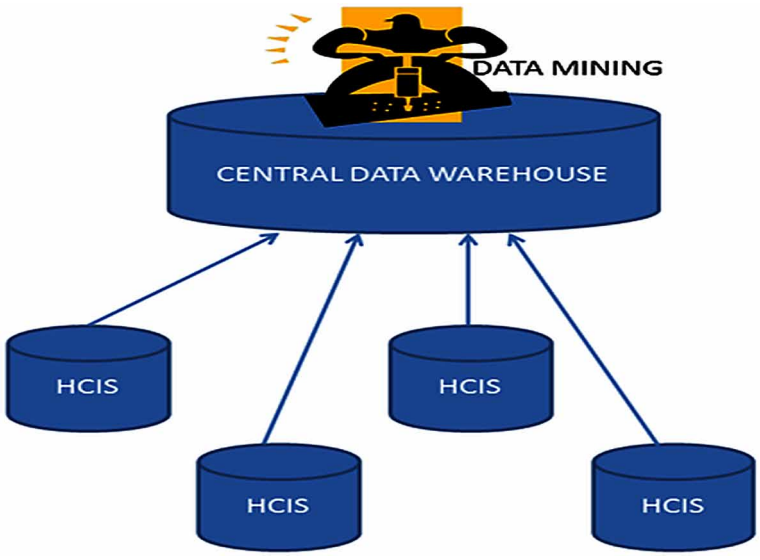
Relation	Notation	Pattern	Name	Inverse Name
Before	A<B	| A | | B |	A before B	B after A
Meets	A m B	| A | | B |	A meets B	B is met by A
Overlaps	A o B	| A | | B |	A overlaps B	B is overlapped by A
Finishes	A f B	| A | | B |	A finishes B	B is finished by A
During	A d B	| A | | B |	A is during B	B contains A
Starts	A s B	| A | | B |	A starts B	B is started by A
Equals	A = B	| A | | B |	A equals B	B equals A

ization of healthcare data, healthcare language problem (called as medical coding systems), electronic communication of the healthcare institutions, collection of patient data scattered to different healthcare institutions, etc. This issue is a challenge to healthcare data mining and will be mentioned in the following sections. There should be a distributed possibility to enable mining scattered healthcare databases and scattered health data of the same patient.

In Figure 3, we see an example of central data mining. All the data in the distributed HCIS databases are transferred to the center data warehouse first, and then they are processed.

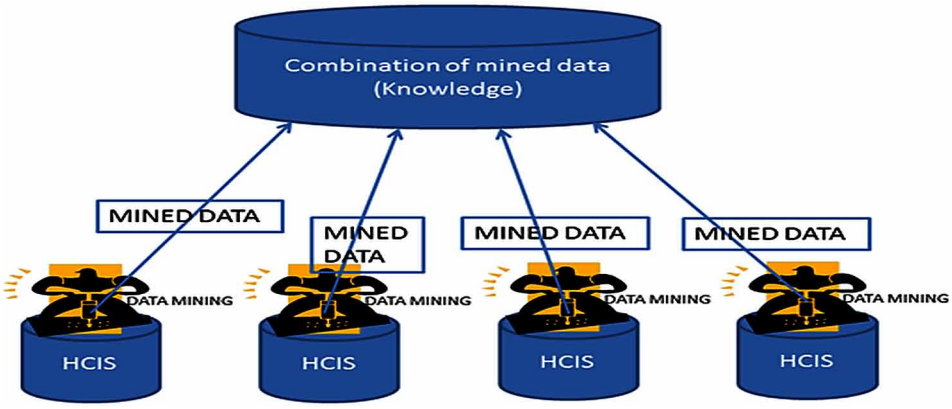
In Figure 4, we see an example of distributed data mining. In this scenario, all the data are processed in the HCIS database cite locally, and the processed results are transferred to the center.

Figure 3. Central data mining



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Distributed data mining



DDM has two possibilities: To process and analyze the data in the distributed centers separately or transfer all data to a center and analyze all centrally. The latter option somehow requires huge resources for the healthcare. There are efforts to have a central database for the patient data but all these efforts have the architecture to have summary data in the center and link them to the original source to avoid the requirement of huge resources and budget. Mining in this central architecture will be nonsense. Then the first option seems more suitable for healthcare

There are efforts to solve DDM in healthcare using agents. The terms agent-based DDM, DDM using intelligent mobile agents, etc. are available in the literature. In such approaches, the idea is deploying a data mining agent to the site and getting analysis from the agent to the center. There is no standard way to accomplish such approaches and it is not easy to perform. There are some studies in the literature having the same idea but proposing different designs.

Challenges for the Use of Data Mining in Healthcare Informatics

In every Introduction to Medical Informatics lecture, the complexity and toughness of the medical environment is emphasized. This "complexity and toughness" is a challenge to every information technology tool and application as well as data mining. In this section of the chapter, the challenges specific to healthcare data mining will be examined.

Heterogeneous Sources and Forms of the Data

In the previous section, scattered nature of the healthcare institutions is mentioned. This scattered structure is a challenge to the healthcare data mining applications and distributed applications are needed to handle this challenge.

The wide range of medical domain also fuels the heterogeneity. There are many different departments of medical domain such as clinics, laboratories, radiology, nuclear medicine, etc. All these departments may have different databases or have a common database if they are the part of the same institution. In either way, they have different type and format of data to mine. For example radiology data have images; laboratory data have number type results whereas clinical data have texts. This "form heterogeneity" is also another challenge for healthcare data mining.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Data Heterogeneity may cause sampling, selection and spectrum bias (Kwiatkowska, Atkins, Ayas, & Ryan, 2007).

Disharmony between Medical and Computer Communities

This challenge is related to the "complexity and toughness" of the medical domain. It is not a specific healthcare data mining challenge but a challenge for all information technology applications. Medical users have tight schedule in dealing with care giving and are not volunteer to use any information technology that costs them additional time. Computer staff urges medical users to enter structured data to have them in computer processable form while medical users want to write down their needs as notes in papers because of the time consumed by dealing with computers.

In addition to this disharmony, the problems of accessibility in hospital exist such as: placing any hardware to sterile areas, entrance of technical staff to these areas for maintenance and fixing can be given as an example.

Legal and Ethical Issues

Privacy and security issues, their protection with the law are another item of challenge that healthcare data mining has to deal with. These issues will be examined in the next section thoroughly. Cios & Moore (2002) describe these challenges as follows:

- **Data Ownership:** Ownership of patient data is not clear. Whether the patients, the physicians, the healthcare institution or the insurance companies own the data collected from patients. The main regulation about the ownership of the data is the Health Insurance Accountability and Portability Act (HIPAA) Privacy Rule.
- **Fear of Lawsuits:** In medical communities, particularly in the Unites States, malpractice and other costly lawsuits is a matter of fear that constitutes a challenge to the application of healthcare data mining.
- **Privacy Issues:** Protecting patient privacy and doctor-patient confidentially constitutes another set of challenges to healthcare data mining. Administrators and researchers should pay utmost attention to privacy and security issues.

Non-Standard Data

This subject deserves being mentioned separately because of its being big challenge for healthcare data mining although it is mentioned in the heterogeneity of data part.

Physician's interpretations in clinical data are written in unstructured form which makes data mining difficult to perform (Milovic & Milovic, 2012). Even the high experienced specialists of the same area cannot agree on common terms that indicate the status of the patient. Different names are used to describe the same disease, different grammatical structures are used to explain the relations between medical entities (Milovic & Milovic, 2012). Although there are considerable efforts to constitute nomenclatures and classifications for structured medical data entry such as International Classification of Diseases (ICD), Systematized Nomenclature of Medicine (SNOMED), International Classification of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nursing Practices (ICNP), still there is enormous size of non-standard data in medical domain. For this reason, data mining researchers have to spend great efforts for text mining applications in healthcare.

Noisy Data

In the previous parts, the tight schedule of healthcare staff is mentioned. This tightness brings many handicaps. Missing, corrupted, inconsistent data are some outcomes. In the time of care giving, the user does not think or care about the usefulness of the data he is entering and he miswrites (if it is obligatory field) or leave it blank. This missing, corrupted, inconsistent data is called as "noisy" in short. Noisy data is a big challenge for data mining. Surely there are data cleaning techniques in data mining to handle noisy data but it causes a bias anyway no matter how well it is handled.

Low Quality of Healthcare Data

In the past, HCISs are used to be designed for financial/billing purposes. This leads to low quality of patient data. Additionally, users' unwillingness to use HCIS contributes to low quality patient data. Unavailability of mobile HCIS usage is another source of low quality data. Especially bed site visits for inpatients, surgery side entries and emergency service operations need mobile data entry to HCIS. If there is no such opportunity, offline entry after a time period elapsed can cause incomplete, inconsistent or missing data in short noisy data.

Domain Participation

Healthcare data mining has different perspectives when compared to data mining in other fields. In the previous section "complexity and toughness" of the medical field is emphasized. This "complexity and toughness" needs domain expertise to perform successful data mining applications. To have both medical and data mining expertise is almost impossible so cooperation and team work is needed for healthcare data mining. This can be possible by either having a domain expert in the research team or convincing medical staff for cooperation.

Complexity of Data Mining Methods

Because each data mining algorithm has its own theoretical design, they require user parameters and end users usually do not have sufficient information about the parameters and their selection (Yoo, Alafaireet, Marinov, Pena-Hernandez, Gopidi, Chang, & Hua, 2012). When it is taken into consideration that data mining results are very sensitive to the parameter, it is a big problem for the success of the data mining task.

Dynamism of the Medical Domain

With the help of the technology, online scientific databases and increasing research funds, medicine is developing rapidly. New devices capturing and storing new forms of healthcare data are invented every day or the current devices are improving by additional facilities. New healthcare diagnosis and prognosis techniques are coming out. In this situation, healthcare data mining becomes a shooter to hit a moving target (Anne, 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Privacy and Security of Patients in Healthcare Data Mining

Privacy and security of patient data became critical with the rapid development of electronic storage of data and HCISs. With the help of electronic facilities, the users get the possibility of accessing and storing vast amounts of data which they never had the ability when they were in paper form. As the data electronically stored in HCIS databases and the size increases, it whets many organizations' appetite. They want to mine and use them for different purposes, legal or illegal. To avoid misusage and confusion, it will be better for readers to give brief descriptions about what is intended with patient privacy and security.

Patient privacy refers to any medical data (medical condition, test result, payment information, etc.) in any form (paper, electronic, etc.) that belongs to a patient. The meaning is about what is protected and who will be permitted to use this information (Upstate Medical University, 2011). The Office for Civil Rights, a department of the United States Department of Health and Human Services, created the Health Insurance Portability and Accountability Act (HIPAA) in 1996 and it is updated regularly. HIPAA has privacy rule for patient privacy. This rule creates a national standard for US to protect medical records of patients. HIPAA privacy rule applies to all forms of patients' protected health information (PHI), whether electronic, written or oral (Center of Medicare & Medicaid Services, 2007).

Data miners should be well aware of what is patient privacy and show utmost effort to comply with the regulations arranging the issue. Otherwise they can harm patients. Data miners can also be subjected to penalties of violating patient privacy.

Security risks are available for all information systems and databases, medical research can add additional risks (Berman, 2002). To provide patient privacy, a privacy review board or institution review board is required to supervise every data mining project in the healthcare institutions (Cooper & Collman, 2005). In the universities, there exists an ethical review board to oversight the ethical issues of any candidate research project. Such a board should be constituted for privacy oversight of the researches related to the patient data.

To provide privacy, three approaches appeared to be common for using personal data in research and secondary analysis of organizations (Lowrance, 2002):

- Use of personal data only with the owner's consent,
- Use of personal data without explicit consent of the owner only in a public interest,
- Making the data anonymous before using.

In many countries, there are laws and regulations against the violation of privacy that apply serious sanctions up to loss of license. Bergman (2002) states that, "Researchers have no automatic right to review patient data. Besides developing strategies for minimizing patient risk, as described herein, investigators should take simple steps to characterize their compliance with human subject's requirements".

Security refers to safeguarding the patient privacy. The meaning is how the information will be protected (Upstate Medical University, 2011). Again HIPAA has a rule for security of the patient data which is called as HIPAA Security Rule.

Healthcare data miners should employ several important security practices such as (Cooper & Collman, 2005):

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- An inclusion of a description of steps to minimize patient risks in the methods sections of research proposals and publication submissions
- For each session and query, good access control and authorization should be used.
- The removal of common identifiers (such as names, addresses) of the data subjects if possible.
- Institution of robust audit practices.
- Training for all principle investigators to reinforce their responsibilities.
- Applications of sanctions for violations of policy and/or procedures.
- Tracking the breaches and sanctions trends over time and usage of these trends in the process of security awareness and training.

Healthcare Informatics Fields of Data Mining

Medical domain is so huge and complex so application of fields of data mining is not an easy task to list. To give an idea about the contribution of data mining to the healthcare, the fundamental fields of data mining applications will be examined in this section.

Hospital Infection Control

Literature says nosocomial infections effect two millions of patients in the USA in each year in 2001 (Gaynes, Richards, Edwards, Emori, Horan, Alonso-Echanove, Fridkin, Peavy, & Tolson, 2001). Association rules can be drawn from culture and patient data obtained from the laboratory database and infection committee can review these rules periodically (Obenshain, 2004).

Standardized Reporting

Hospitals report many incidents and epidemic diseases to the government. Cluster and association analyses can show how risk factors are reported with the use of International Classification of Diseases and by reconstructing patient profiles (Obenshain, 2004).

Identifying High Risk Patients

Data mining applications can report patients that have high risks. This can aid caregivers to take some precautions and save lives (Obenshain, 2004).

Outlier Detection

Outlier Detection, can be also named as Anomaly detection, provides opportunity for identifying rare events in large datasets (Jacob & Ramani, 2012). Detection of extreme observations may prevent incorrect data and human error while at the same time presence of Outliers could lead to novel insights in clinical knowledge discovery (Jacob & Ramani, 2012).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fraud and Abuse Detection

Healthcare is prone to fraud and abuse because of the huge amount of money it creates. Both healthcare providers and clients can commit fraud. In healthcare, the most common forms of provider fraud include (Patel, Hsu, & Lee, 2008):

- Billing for services not furnished,
- Misrepresenting the diagnosis to justify payment,
- Soliciting, offering or receiving a kickback,
- Unbundling or "exploding" charges,
- Falsifying certificates of medical necessity, plans of treatment and medical records to justify payment,
- Billing for a service not furnished as billed.

Client abuses can be card-sharing, acting in collusion with a provider for kickbacks, etc. (Patel, Hsu, & Lee, 2008).

Especially medical insurance companies employ data mining applications to detect whether healthcare claims are potentially fraudulent or erroneous, to reduce their losses related to fraud and abuse.

Treatment Effectiveness

Data mining can help about which courses of action prove effective by comparing and contrasting causes, symptoms and courses of treatments (Koh & Tan, 2011). Data mining also helps evolve standardized treatments for specific diseases (Koh & Tan, 2011). Also applications to facilitate early diagnosis, non-invasive diagnosis, determination of adverse drug effects can be made.

Healthcare Management

This section is examined thoroughly in upcoming sections.

Customer Relationship Management (CRM)

In the previous parts it is mentioned that healthcare institutions have to increase the quality. This increase comprises increasing the customer satisfaction as well as the services given. The customers could be patients, pharmacists, physicians or clinics. Detection of usage and purchase patterns and the eventual satisfaction they result in can be used to improve overall customer satisfaction, in many cases, purchasing and usage behavior prediction aid in designing proactive initiatives to reduce the costs and increase the customer satisfaction (Desikan, Hsu, & Srivastava, 2011).

Data mining applications can help healthcare to determine the preferences, usage patterns, current and future needs of individuals, whether a patient is likely to comply with prescribed treatment or whether preventive care is likely to produce a significant reduction in future utilization (Koh & Tan, 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Reduction of Medical Errors

Medical errors are an important issue in healthcare. According to the report of Institute of Medicine (IOM) at least 44,000, may be as many as 98,000 patents die in the hospitals each year because of medical errors (Duan, Street, & Xu, 2011). Another study states that 195,000 people in the U.S. died due to potentially preventable, in-hospital medical errors out of 37 million patient records (Amin, Agarwal, & Beg, 2013). In addition to the loss of human lives, medical errors have a huge cost of both compensation and loss of reputation. It is seen that medical errors are harmful and costly. This is a giant problem to tackle with. One of the most important application fields of data mining is medical errors. By means of decision support facility, it can help reduce medical errors.

Data Mining for Healthcare Management

Healthcare management is a separate area of science in itself because of its importance and complexity. There are departments in universities about this issue, Master of Science and Ph.D. programs are available on this topic. Naturally there are job positions in the healthcare institutions as "Healthcare Manager". All the educations, researches, positions in healthcare management have one big goal: Managing the healthcare efficiently. Tasks of the healthcare managers can be described as optimization of clinical processes in medical and administrative quality terms as well as the cost/benefit relation (Milovic & Milovic, 2012).

Healthcare managers make very critical decisions about amount of supplies, number of employees, drug needs, hardware and software management, vacant beds required, maintenance needs,etc. These decisions must be on time, true and well supported. To do this, like all management branches, healthcare management needs the effective use of informatics maybe more than any other management field. They need approximate and healthy predictions.

Data mining applications are useful for better identification of and tracking chronic disease states and high-risk patients, for designing appropriate interventions and reducing the number of hospital admissions; claims would help healthcare management (Koh & Tan, 2011). By means of data mining, comparisons across healthcare groups of things such as practice patterns, resource utilization, length of stay, and costs of different hospitals be possible (Johnson, 2001), which are valuable information for the healthcare management.

Data mining can help healthcare managers to (Milovic & Milovic, 2012):

- Discover new hypothesis for indexes of quality for data, standards, plans and treatments;
- Check if the given quality for data, standards, plans and treatments indexes are still valid;
- Improve, strengthen and adjust the quality indexes for data, standards, plans and treatments;

Resource allocation is another field that data mining can help healthcare management. By detecting high risk areas, it should also predict the requirement and usage of various resources accordingly (Sharma & Speights, 2014).

One of the major problems in healthcare is stated as measuring the flow of patients through hospitals along with other healthcare facilities that are being used (Sharma & Mansotra, 2014). If healthcare managers can predict the length of stay of inpatients correctly, planning and management of hospital resources can be easier (Sharma & Mansotra, 2014).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Healthcare management is not limited to the management of a healthcare institution. It also encompass the public health services given, directed, supervised by governmental organizations. They need multiple sourced data mining applications to manage the public health. In this point, DDM applications are needed for this kind of healthcare management. These applications help managers manage some critical risks such as disease outbreak, epidemic diseases, some trends such as cost of healthcare, mortality rates, birth rates, appearance frequencies of the diseases, their sub information about the cities, ages, etc.

These examples about how data mining may help healthcare management can easily be augmented. The ability of data mining to associate rules and uncover hidden relations provides invaluable support to healthcare management both in organization level and governmental level. Key questions of the process of healthcare quality management are quality of data, standards, plans and treatments all of which can be acquired by the utilization of available data at hand.

FUTURE RESEARCH DIRECTIONS

Healthcare domain is dynamic and we named this issue as a challenge to the data mining. It is also a forcing motivation to the data mining applications in healthcare. This dynamism gives way to new horizons and more data mining applications will be employed to discover new patterns and associations.

In the view of the subjects examined in this chapter, future data mining studies seem to take place, not limited to but in considerable weight, in distributed data mining applications. Nowadays, hospital chains are emerging. All over a country or city, there exist a number of hospitals belonging to an organization. They have distributed solutions as well as central ones. Governmental prediction and management issues are another source of need for DDM applications. In short, needs for DDM applications are growing which is making the DDM a future progress area candidate.

Text mining algorithms are another candidate area of progress. Although the idea of HCIS force users to structured data entry, healthcare is not ready for a fully structured data entry and there are doubts if it can be provided fully. Radiology reports, surgery reports, Nuclear Medicine reports, anamnesis forms and more examples about healthcare are in fully text data form. For mining these data, text mining methodologies will get great efforts most probably.

The low quality of the medical data is mentioned throughout the chapter. In academic literature, these kinds of data are called as dirty and noisy. Although there are rigorous data mining algorithms available that deal with these problems, and there is a preprocessing phase of data mining, healthcare data deserves more research and development in this issue because of its bad reputation and nature.

While developing and improving data mining, patient privacy and security should be the first priority of concern. Health data is the most private data of a person. There may be so sensitive information that can make a person ashamed and upset. There may be some details even the person himself wants to forget. Because of the sensitivity of the data, there should be considerable future work for protecting the security and privacy of EPHI both by private sector and governmental bodies.

CONCLUSION

In the era of information and technology, data is too valuable. Every possibility that lit a light about what future can look like is very welcomed by the organizations. When it comes to healthcare, this importance

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

increases exponentially, because human life is the main subject of the sector. By the help of the improvements in science, the deadly diseases are so trivial now that we do not fear anymore.

Data and its processing is a catalyzer to the improvement in all areas of science. In this era, information and knowledge is the power. Data is all over, the magic is to convert it to information and knowledge, to get the power.

Complexity and toughness of the medical domain makes use of information technologies mandatory. Its scattered nature, text loaded data forms, having the subject of human life and the pressure to decrease costs push healthcare managers to use data effectively and convert it to the information and knowledge. Accomplishment of these tasks is possible with the effective use of data mining.

ACKNOWLEDGMENT

Part of this study was presented at the 2nd International Conference on e-Health and Telemedicine ICEHTM 2014, on 22-24 May 2014, Istanbul/Turkey.

REFERENCES

Allen, J. F. (1984). Towards a general theory of action and time. *Artificial Intelligence*, *23*(2), 123–154. doi:10.1016/0004-3702(84)90008-0

American Medical Association. (2010). *Getting the most for our health care dollars: Medical liability reform.* Chicago, IL: American Medical Association. Retrieved on 30 December 2015 from http://www.allhealth.org/briefingmaterials/AMASharedDecisionMaking-1936.pdf

Amin, S. U., Agarwal, K., & Beg, R. (2013). Data mining in clinical decision support systems for diagnosis, prediction and treatment of heart disease. *International Journal of Advanced Research in Computer Engineering & Technology*, *2*(1), 218–223.

Anne. (2011). *Challenges for process mining in medicine*. Retrieved on 30 August 2013 from http://www.techcrunch.com

Batal, I., Valizadegan, H., Cooper, G. F., & Hauskrecht, M. (2012). A Temporal pattern mining approach for classifying electronic health record data. *ACM Transactions on Intelligent Systems and Technology*, *4*(4), 63. PMID:25309815

Berman, J. J. (2002). Confidentiality issues for medical data miners. *Artificial Intelligence in Medicine*, *26*(1-2), 25–36. doi:10.1016/S0933-3657(02)00050-7 PMID:12234715

Bramer, M. (2013). *Principles of data mining*. London: Springer-Verlag. doi:10.1007/978-1-4471-4884-5

Bruno, G., & Garza, P. (2012). Temporal pattern mining for medical applications. In *Data Mining: Foundations and Intelligent Paradigms* (pp. 9–18). Berlin: Springer. doi:10.1007/978-3-642-23151-3_2

Canlas, R. D., Jr. (2009). *Data Mining in Healthcare: Current Applications and Issues* (Unpublished master dissertation). Carnegie Mellon University, Australia.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Center of Medicare & Medicaid Services. (2007). *Security 101 for covered entities*. Retrieved on June 10, 2015 from: http://www.hhs.gov/ocr/privacy/hipaa/administrative/securityrule/security101.pdf

Chen, H., Fuller, S. S., Friedman, C., & Hersh, W. (2005). Knowledge management, data mining, and text mining in medical informatics. In Medical Informatics (pp. 3-33). Springer US. doi:10.1007/0-387-25739-X_1

Cios, K. J., & Moore, G. W. (2002). Uniqueness of medical data mining. *Artificial Intelligence in Medicine*, *26*(1), 1–24. doi:10.1016/S0933-3657(02)00049-0 PMID:12234714

Concaro, S., Sacchi, L., Cerra, C., Fratino, P., & Bellazzi, R. (2011). Mining health care administrative data with temporal association rules on hybrid events. *Methods of Information in Medicine*, *50*(2), 166–179. doi:10.3414/ME10-01-0036 PMID:21170468

Cooper, T., & Collman, J. (2005). Managing information security and privacy in healthcare data mining. In Medical Informatics (pp. 95-137). Springer US. doi:10.1007/0-387-25739-X_4

Desikan, P., Hsu, S., & Srivastava, J. (2011, April). *Data mining for health care management*. Paper presented at 2011 SIAM International Conference on Data mining, Mesa, AZ.

Duan, L., Street, W. N., & Xu, E. (2011). Healthcare information systems: Data mining methods in the creation of a clinical recommender system. *Enterprise Information Systems*, *5*(2), 169–181. doi:10.1080/17517575.2010.541287

Gaynes, R., Richards, C., Edwards, J., Emori, T. G., Horan, T., Alonso-Echanove, J., ... Tolson, J. (2001). Feeding back surveillance data to prevent hospital-acquired infections. *Emerging Infectious Diseases*, *7*(2), 295–298. doi:10.3201/eid0702.010230 PMID:11294727

Hays, T. (2012, December). *Medical data mining*. Speech made at The National Institute of Standards and Technology (NIST) Symposium. Retrieved on 30 December 2015 from http://www.nist.gov/healthcare/upload/Hays-Medical-Data-Mining-slides-for-web.pdf

Henriques, R., Pina, S., & Antunes, C. (2013). Temporal mining of integrated healthcare data: Methods, revealings and implications. *SDM IW on Data Mining for Medicine and Healthcare*, 52-60.

Jacob, S. G., & Ramani, R. G. (2012). Data mining in clinical data sets: A review. *International Journal of Applied Information Systems*, *4*(6), 15–26. doi:10.5120/ijais12-450774

Jain, N., & Srivastava, V. (2013). Data mining techniques: A survey paper. IJRET. *International Journal of Research in Engineering and Technology*, *2*(11), 1163–2319.

Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs*. London: Springer. doi:10.1007/978-0-387-68282-2

Jiawei, H., & Kamber, M. (2001). Data mining: Concepts and techniques. San Francisco, CA: Morgan Kaufmann.

Johnson, D. E. (2001). Web-based data analysis tools help providers, MCOs contain costs. *Health Care Strategic Management*, *19*(4), 16. PMID:11329971

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kantardzic, M. (2011). *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons. doi:10.1002/9781118029145

Koh, H. C., & Tan, G. (2011). Data mining applications in healthcare. *Journal of Healthcare Information Management*, *19*(2), 65. PMID:15869215

Kolçe, E., & Frasheri, N. (2012, September). *A literature review of data mining techniques used in healthcare databases*. Paper presented at the ICT innovations 2012, Ohrid, Macadonia.

Kwiatkowska, M., Atkins, M. S., Ayas, N. T., & Ryan, C. F. (2007). Knowledge-based data analysis: First step toward the creation of clinical prediction rules using a new typicality measure. Information Technology in Biomedicine. *IEEE Transactions on*, *11*(6), 651–660.

Laboratory for Biomedical Informatics. (2015). Retrieved on September 11, 2015, from http://www.labmedinfo.org/translational/temporal

Lin, W., Orgun, M. A., & Williams, G. J. (2002). Mining temporal patterns from health care data. In *Data Warehousing and Knowledge Discovery* (pp. 222–231). Springer Berlin Heidelberg. doi:10.1007/3-540-46145-0_22

Lowrance, W. W. (2002). Learning from experience: Privacy and the secondary use of data in health research. *The Journal of Biolaw & Business*, *6*(4), 30–60. PMID:15072055

Milovic, B., & Milovic, M. (2012). Prediction and decision making in Health Care using Data Mining. *International Journal of Public Health Science*, *1*(2), 69–78.

Mishra, P., Padhy, N., & Panigrahi, R. (2013). The survey of data mining applications and feature scope. *Asian Journal of Computer Science & Information Technology*, *2*(4), 43–58.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(3), 247–259. doi:10.1016/j.isprsjprs.2010.11.001

Obenshain, M. K. (2004). Application of data mining techniques to healthcare data. *Infection Control and Hospital Epidemiology*, *25*(8), 690–695. doi:10.1086/502460 PMID:15357163

Patel, D., Hsu, W., & Lee, M. L. (2008, June). Mining relationships among interval-based events for classification. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 393-404). ACM. 10.1145/1376616.1376658

Sarvestani, A. S., Safavi, A. A., Parandeh, N. M., & Salehi, M. (2010, October). Predicting breast cancer survivability using data mining techniques. In *Software Technology and Engineering (ICSTE), 2010 2nd International Conference on* (Vol. 2, pp. V2-227). IEEE. 10.1109/ICSTE.2010.5608818

Sharma, A., & Mansotra, V. (2014, March). Emerging applications of data mining for healthcare management-A critical review. In *Computing for Sustainable Global Development (INDIACom), 2014 International Conference on* (pp. 377-382). IEEE. 10.1109/IndiaCom.2014.6828163

Speights, K. (2013). 7 Countries with the highest health care costs. *Daily Finance*. Retrieved on September 14 2015 from http://www.dailyfinance.com/2013/06/15/7-countries-with-the-highest-health-care-costs/

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sun, J., Zhu, X., Zhang, C., & Fang, Y. (2012). Security and privacy for mobile health-care (m-Health) systems. In S. K. Das, K. Kant, & N. Zhang (Eds.), *Handbook on Securing Cyber-Physical Critical Infrastructure* (pp. 677–704). Elsevier. doi:10.1016/B978-0-12-415815-3.00027-3

Sundar, N. A., Latha, P. P., & Chandra, M. R. (2012). Performance analysis of classification data mining techniques over heart disease database. *International Journal of Engineering Science & Advanced Technology*.

Takabayashi, K., Ho, T. B., Yokoi, H., Nguyen, T. D., Kawasaki, S., Le, S. Q., ... Yokosuka, O. (2007). Temporal abstraction and data mining with visualization of laboratory data. *Studies in Health Technology and Informatics*, *129*(2), 1304. PMID:17911925

Upstate Medical University. (2011). *Protecting patient confidentiality and security*. Retrieved on September 1 2015, from www.upstate.edu/forms/documents/F84037.pdf

Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *Knowledge and Data Engineering*. *IEEE Transactions on*, *26*(1), 97–107.

Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, J. F., & Hua, L. (2012). Data mining in healthcare and biomedicine: A survey of the literature. *Journal of Medical Systems*, *36*(4), 2431–2448. doi:10.100710916-011-9710-5 PMID:21537851

ADDITIONAL READING

Brameier, M., & Banzhaf, W. (2001). A comparison of linear genetic programming and neural networks in medical data mining. Evolutionary Computation. *IEEE Transactions on*, *5*(1), 17–26.

Chen, J. Y., & Lonardi, S. (Eds.). (2009). *Biological data mining*. London: CRC Press.

Cios, K. J., & Kacprzyk, J. (2001). *Medical data mining and knowledge discovery*. New York, NY: Physica-Verlag.

Dua, S., & Acharya, U. R. (2011). *Data mining in biomedical imaging, signaling, and systems*. London: Auerbach Publications.

Epstein, I., & Blumenfield, S. (2001). *Clinical data-mining in practice-based research: Social work in hospital settings*. New York: Routledge.

Evfimievski, A., & Grandison, T. (2009). Privacy-preserving data mining. Handbook of Research on Innovations in Database Technologies and Applications: Current and Future Trends, 527-536. doi:10.4018/978-1-60566-242-8.ch056

Fu, T. C. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, *24*(1), 164–181. doi:10.1016/j.engappai.2010.09.007

Gada, F. V., Yadav, K. R., & Shah, R. D. (2013). Data Mining in Medical Application for better Detection of Disease. *IJCER*, *2*(2), 173–176.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Giles, R., Epstein, I., & Vertigan, A. (Eds.). (2011). *Clinical data mining in an allied health organisation: A real world experience*. Sydney: Sydney University Press.

Iakovidis, D., & Smailis, C. (2012). A semantic model for multimodal data mining in healthcare information systems. *Studies in Health Technology and Informatics*, *180*, 574–578. PMID:22874256

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: Towards better research applications and clinical care. *Nature Reviews. Genetics*, *13*(6), 395–405. doi:10.1038/nrg3208 PMID:22549152

Khan, A., Doucette, J., Jin, C., Fu, L., & Cohen, R. (2011, January). An ontological approach to data mining for emergency medicine. In 2011 Northeast Decision Sciences Institute Conference Proceedings 40th Annual Meeting. Montreal, Quebec, Canada (pp. 578-594).

Lavrač, N., & Zupan, B. (2010). Data mining in medicine. In Data Mining and Knowledge Discovery Handbook (pp. 1111-1136). Springer US.

Maulik, U., Bandyopadhyay, S., & Mukhopadhyay, A. (2011). Multiobjective genetic algorithms for clustering: Applications in data mining and bioinformatics. Springer USA.

Prather, J. C., Lobach, D. F., Goodwin, L. K., Hales, J. W., Hage, M. L., & Hammond, W. E. (1997). Medical data mining: Knowledge discovery in a clinical data warehouse. In *Proceedings of the AMIA Annual Fall Symposium* (p. 101). American Medical Informatics Association.

Raipure, G. S., & Gupta, S. R. (2013). Decision support system in medical science using OLAP & data mining. *International Journal Of Computer Science And Applications*, *6*(2), 248–254.

Ramakrishnan, N., Hanauer, D., & Keller, B. (2010). Mining electronic health records. *Computer*, *43*(10), 77–81. doi:10.1109/MC.2010.292

Rao, R. B., Sandilya, S., Scherpbier, H. J., & Warrick, T. R. (2012). U.S. Patent No. 8,214,224. Washington, DC: U.S. Patent and Trademark Office. Retrieved on Jan 01, 2016 from https://www.google.com/patents/US8214224

Smith, M. R., Wang, X., & Rangayyan, R. M. (2009). Evaluation of the sensitivity of a medical data-mining application to the number of elements in small databases. *Biomedical Signal Processing and Control*, *4*(3), 262–268. doi:10.1016/j.bspc.2009.04.001

Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computers and Applications*, *17*(8), 43–48. doi:10.5120/2237-2860

Tuffery, S. (2011). *Data mining and statistics for decision making*. John Wiley & Sons. doi:10.1002/9780470979174

This research was previously published in Intelligent Multidimensional Data Clustering and Analysis edited by Indrajit Pan, Sourav De, Siddhartha Bhattacharyya, and Paramartha Dutta; pages 135-158; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 4
Population Health Management and the Science of Individuality

Anastasius Moumtzoglou
Hellenic Society for Quality and Safety in Healthcare, Greece, & P. & A. Kyriakou Children's Hospital, Greece

Abraham Pouliakis
National and Kapodistrian University of Athens, Greece

ABSTRACT

This article espouses that population health management (PHM) has been a discipline which studies and facilitates care delivery across a group of individuals or the general population. In the context of population health management, the life science industry has had no motivation to design drugs or devices that are only effective for a distinct segment of the population. The major outgrowth of the science of individuality, as well as the rising 'wiki medicine', fully recognizes the uniqueness of the individual. Cloud computing, Big Data and M-Health technologies offer the resources to deal with the shortcomings of the population health management approach, as they facilitate the propagation of the science of individuality.

INTRODUCTION

Public health' connotes a relatively narrow field with activities carried out by agencies granted with official functions. 'Population health' a term with broader content, is related to a field relevant to the study of several important factors for health. As a result, it involves many terms, such as outcomes, disparities, determinants, and risk factors (Kindig, 2007).

Notwithstanding the term 'population health' combines the concepts of both health and population, every term has an essential meaning of its own. The population is related to a group of persons being organized into numerous different units of analysis. Similarly, the term health was defined negatively, i.e., the absence of disease. Nowadays, the modern understanding stresses the positive aspects as well, and health is considered to be related to all life issues. Summarizing, today it is debated whether population health and public health are different or identical. Nevertheless, population health is defined as health outcomes and their distribution in a population (Kindig, 2007).

DOI: 10.4018/978-1-7998-1204-3.ch004

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Population health management (PHM) has been defined as 'the technical field of endeavor which utilizes a variety of individual, organizational and cultural interventions to help improve the morbidity patterns (i.e., the illness and injury burden) and the health care use behavior of defined populations (Hillman, 2002). It is differentiated from disease management because it includes (Howe & Spence, 2004):

- More chronic conditions and diseases.
- Uses a single point of contact and coordination and predictive modeling across multiple clinical conditions.

Moreover, PHM is considered a broader term than disease management, as it includes (Coughlin, Pope, & Leedle(Jr), 2006):

- Intensive care management for individuals at the highest level of risk.
- Personal health management for those at lower levels of predicted health risk.

At the provider level, there are three highlighted components (Care Continuum Alliance, 2012):

- The leadership and the central care delivery role of the primary care physician.
- The critical importance of patient activation.
- The capacity expansion of care coordination.

In this context, to successfully achieve all of these requirements, an organization should provide proactive, preventive and chronic care services to all managed patients. Additionally, this should take place both during encounters of patients with the healthcare system and in between. Therefore, providers should maintain regular contact with their patients and support them in the management of their health. Additionally, providers must manage patients at high risk, to prevent the deterioration of their health and avoid the development of complications. Finally, evidence-based protocols for the diagnosis and treatment of patients, consistently and cost-effectively, are also required if for a provider-based PHM approach is followed.

The Federal Agency for Healthcare Research and Quality (AHRQ) developed the concept of 'practice-based population health' (PBPH) and defined it as "an approach to care that uses information on a group of patients within a primary care practice or group of practices to improve the care and clinical outcomes of patients within that practice.' (Cusack, Knudson, Kronstadt, Singer, & Brown, 2010).

Obviously, the implementation of health IT is among the essential components and requirements for planning and implementing PHM. The adoption of Electronic Health Records is just the first step toward the creation of the requisite infrastructure. However, a wide range of other IT applications to automate PHM, track results and engage the patients in their health care is required. Additionally, IT systems should repeatedly be evaluated because of rapid technological changes, new government regulations and new approaches to patient management (Moumtzoglou & Pouliakis, 2015).

AHRQ recommends health IT tools for the stratification and monitoring of populations, as follows:

- Target patients being in greatest need of health services by stratifying the patients and narrowing subpopulations.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Make patient-related data actionable by generating alerts to the patients.
- Make data actionable by generating alerts towards health providers, about patient care needs.

Conclusively, population health management is designed around the collective. However, nowadays, to truly drive high-quality care at lower cost, these endeavors must work for the individual, align with their personal goals and intermix with their reality. In this context, the objective of the article is twofold. First, it will deal with the greatest misconception about population health, that is, population health is not about the population. Ultimately, it is about the individual. Secondly, it will show how to connect the dots from the science of individuality to improving health on a massive scale - a bottom-up approach that uses the technologies of cloud computing, Big Data and M-Health.

THE SCIENCE OF INDIVIDUALITY

Unquestionably, the population health management concept requires extensive data, which currently flows from the beneficiaries towards data centers. Such data include a multitude of health and well-being information collected manually or automatically. Moreover, algorithms are required to process such information and alert when health-related actions should be taken. On the other hand, nowadays, there are technologies which provide both the data storage and processing capacity while the mobile device (especially the mobile phone), as well as inter-connectable devices (health gadgets), may serve as data acquisitions points on the individual level.

However, there is a major deficiency of contemporary medicine which is reflected by the use of experts to make recommendations or guidelines for a large proportion of decisions for which no or minimal data exists. These guidelines are published in major specialty journals which have a pronounced impact, as they are believed to represent the standard of care. Overall, there is too much literature which is evidenced by the statistic that only 0.5 percent of the 38 million published papers are cited more than two hundred times, and half were never cited. Moreover, when pooled analyses of prior studies are published, many relevant papers are excluded (Topol, 2012).

In fact, this should be considered ‘eminence-based medicine.’ (Topol, 2012).

Better studies might be part of the solution. However, as we can accrue more meaningful data and information for individuals, the hope is that we can override our dependence on such recommendations. We need evidence which is built not on populations but individuals.

Fortunately, our capacity to get just that information is rapidly arising. An era which is characterized by the right drug and dose, at the right cost and timing has begun (Topol, 2012). Medicine for the common good is not sufficient. A rebooted life science, instead of a mass-population-directed strategy, would leverage the science of individuality, getting the relevant digital readout from an individual to shape a therapy. We now, for the first time, have the devices to, promote a level of prescription precision we have never seen before (Topol, 2012).

The era of individualized medicine ultimately promises to do away with terms like ‘cryptogenic’ and ‘essential,’ fully recognizing the uniqueness of the individual who needs to be seen and treated with utter respect for his or her individuality. It will not be long until digitizing a person unlocks the cause for what is wrong, creating valuable knowledge that can save a life or markedly improve the quality of life. That is a major outgrowth of the science of individuality.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In this context, the entire classification system of medical conditions and diagnoses is about to be rewritten. Instead of our current reductionist model for which individuals are unwisely assigned to such categories as one of two types of diabetes or cancer of a particular organ, the science of individuality will promote a new molecular taxonomy that invokes the main biologic basis, regarding genes or pathways.

THE TECHNOLOGIES OF INDIVIDUALITY

The data for how many patients search health websites are staggering (Topol, 2012). Although this trend represents an outlier in many respects and an extreme of patient activism, these cases also demonstrate that the current level of being informed is not adequate. Furthermore, they capture a sense of inspiration and independence, the hunt for innovation, and the primacy of the individual. In essence, they are the precursors for the next phase of medicine, in which powerful digital tools will provide data that was heretofore unavailable.

Whether it is sequencing the genome of cancer tissue to determine a specific mutation or being able to anticipate the likelihood of a fatal disease before it has ever manifested, our capabilities greatly expand based on the following technologies.

Cloud Technologies

Cloud computing has emerged rapidly as an exciting new paradigm that offers a challenging model of computing and services. Leveraging cloud computing technology, bioinformatics tools, can be made available as services to anyone, anywhere, and through any device. The use of large bio-datasets, the highly demanding algorithms, and the hardware for sudden computational resources makes large-scale bio-data analysis an attractive test case for cloud computing" (Hsu, Lin, Ouyang, & Guo, 2013). Cloud computing is a concept related to numerous computers that use Internet infrastructure to communicate producing shared resources such as software, hardware, and storage. Nowadays, cloud computing is not widely used for the various tasks related to medicine; however, there are numerous fields that it could be employed; thus, cloud computing is an emerging field having the potential to change the everyday practice of medicine.

Cloud systems are categorized into three different groups according to the offered service type:

- Infrastructure as a Service (IaaS), which means offering hardware, storage and physical devices over the Internet.
- Software as a Service (SaaS), which means offering software and hosted applications over the Internet.
- Platform as a Service (PaaS), which means offering the capability to deploy applications created using programming languages, libraries, services, and tools owned and supported by the provider.

For cloud computing, the user/consumer does not manage or control the underlying cloud infrastructure but has control over the deployed applications as well as the offered services (i.e. storage, networking and processing power). Clouds, according to their hosting, may be public, private, hybrid, and community, especially adapted to the medical field:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Public Clouds are for general use. The cloud owners are responsible for information hosting; public clouds are rarely utilized in the area of medicine and case of use, data are encrypted.
- Private Clouds are only for in-hospital use and are dealing with confidential patient data. The owners or the hospital premises are responsible for information hosting.
- Hybrid Clouds are hosting non-confidential information on public Clouds and confidential information in a private domain.
- Community Clouds are hosting information among members of the same community. Laboratories and hospitals can create a community Cloud while sharing the same infrastructure and software.

There are three fundamental aspects of Cloud-based services (Soman, 2011):

Browser-Based Applications

Browser-based applications are different from native applications, which run on personal computers. Native Applications are those that run entirely on the computer on which they are invoked. A native application uses only the computing resources and storage of the computer which was used to launch the application. Conversely, browser-based applications are applications that run from within Internet Browsers; however, the consumed resources are located on one or more servers. They are conceptually similar to Client Server applications, which require a powerful central computer, but there are two distinct characteristics:

- Browser-based applications can only be invoked from inside a standard browser.
- The Client and the Server are not (necessarily) a part of a Local Area Network or a remote (however accessible) network.

Optimal Utilization of Server Resources

Cloud Service Providers see variable levels of utilization of their services. This variability has two components. Firstly, the number of concurrent users accessing the service is different at different times. Secondly, the processing and storage of resources required by any particular user are never constant.

Technologies for scaling and resource partitioning are at the heart of Cloud Services because without these technologies, the Cloud business model would not be feasible. Load Balancing denotes a mechanism for balanced or even distribution of load across multiple physical resources, such as servers, storage devices or communication links.

A server load balancer is responsible for balancing the load of incoming requests among an array of physical servers. Virtualization allows a single physical resource to be used like many resources or multiple physical resources to be used as a single resource. There are several types of virtualization, depending on the nature of the physical resource that is 'virtualized.' In this context, server virtualization implies that a single physical server is converted into several logical servers, storage virtualization a single storage device that is converted into multiple logical storage devices, or vice-versa, and network virtualization allows the user to split a single communication link into separate channels assigned to different processes.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Finally, multi-tenancy is the principal of architecting applications such that a single instance of the application running on the server is visible as separate cases of the same application for each of the clients connected to it.

Data Centers

The data center is where the user's data is processed and stored. Some of the design considerations for data centers are as below (Moumtzoglou, 2014; Soman, 2011):

- Data centers are typically located in places where the risk of natural disasters and human-made disasters.
- Buildings used to house data centers typically do not accommodate other offices and businesses. The area around the data center is well-lit and easy to monitor. It is easy to control access to the premises where the data center is housed.
- Data centers have areas with raised flooring.
- Data centers have to air-condition.
- Data centers have different levels of physical security.
- Data centers have assured stable electric supply.
- Data centers have state of the art fire prevention and fire protection systems.

From a networking perspective (Moumtzoglou, 2014; Soman, 2011):

- Data centers have high-bandwidth data pipes with various levels of redundancy.
- Data centers have all the elements installed that are required to run Intranet services continually.
- Data centers achieve the most advanced network security tools including intrusion detection, firewalls, antivirus systems, and systems that guard against spamming, Denial of Service Attacks, and all known malware.

Lastly (Moumtzoglou, 2014; Soman, 2011):

- Data centers employ qualified system administrators.
- Data centers support systems and processes in place to keep smooth functioning 24x7 and throughout the year.
- Data centers have a disaster recovery plan.

CLOUD-BASED SYSTEMS FOR HEALTHCARE IT

The healthcare industry is different from most other industries. It is very delicate, and there are several entities that have to deal with healthcare data. That includes regulatory bodies, care providers, hospital administration staff, payers, labs, and patients themselves. The key differences between the healthcare industry and other industries, in which Cloud-based systems have been used, include (Soman, 2011):

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **High-Risk Industry:** The healthcare industry impacts individuals to a greater extent than any other industry.
- **Highly Regulated:** Numerous regulations govern the provision of healthcare.
- **Multiple Stakeholders:** There is a large number of interested parties in the healthcare industry. That includes the patients, the physicians, the nurse practitioners, the hospitals/clinics, their administrative staff members, the payers/ insurance companies, employers, pharmaceutical companies, technology vendors, device manufacturers, government bodies, and so on.
- **Slow Pace of Adoption:** The healthcare industry has traditionally been slow to adopt IT, especially the smaller clinics and individual physician practices.
- **Small and Large Providers:** The small providers typically do not have the resources to evaluate or experiment with advanced IT systems. On the other hand, there are several large multi-specialty hospitals whose IT requirements are distinctly different from those of office practitioners.
- **Long-Term Relationships:** Changing IT systems often is not easy, and there is, therefore, an assumption that any product adopted would be advantageous to use for several years. This hypothesis does not follow the need for both the IT systems and individuals to adapt.

Given all of these differences, it is not clear that Cloud-based systems would work as well in the healthcare industry as they have in other industry verticals unless Cloud-based healthcare solutions address these specific attributes of the industry.

From a clinical perspective, an EMR (electronic medical record) leads to improvement in the quality of care provided to patients. Cloud-based EMRs are, therefore, a key component of any Cloud-based healthcare offering. An EMR is a computer-based documentation of a patient's medical history including illnesses, examinations, reports, diagnosis and the care administered. Electronic Medical Records are created and maintained by care providers when a patient avails service from the vendor.

The following key pieces of medical information are a part of virtually all EMR systems (Soman, 2011):

- Patient Identification information
- Information about Patient Allergies and Habits
- Immunization Record
- Medical History
- Benefits of EMR

There are significant advantages to using EMR when compared to the traditional paper-based files and charts (Soman, 2011):

- Storage
- Usage in Emergency
- Duplication and Transportation costs
- Uniformity
- Risks
- Audit/Reporting
- Integration
- Efficiency

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cloud-based EMR systems have all the features and capabilities of in-house EMR systems. Many Cloud-based systems have rich and especially easy-to-use interfaces along with the ability to customize the fields in the records. Cloud-based EMR systems make it even easier to access or transfer medical information to any point, for example, to the point of care in an emergency (Witzke & Specht, 2017). Cloud-based EMR systems also make it simpler to provide access to the records of the patient via patient portals.

A Cloud-based medical practice management application enables a healthcare provider to manage and streamline business management tasks and workflows. A 'workflow' is a series of steps followed to achieve a specific purpose. Provision of healthcare involves the implementation of various workflows within the practice, mirroring the movement of the patient within the system. Medical Practice Management relates to the operation of various workflows within the practice while Medical Practice Management Systems are the key to ensuring that all transactions within the practice run smoothly and efficiently.

Cloud-based Practice Management software supports features such as (Soman, 2011):

- Appointment scheduling of patients
- Eligibility and authorization
- Physician scheduling, scheduling surgery or other procedures for patients
- Tracking patient referrals
- Patient account management
- Managing patients as they move from admission to discharge including information on hospital rounds
- Managing patient recall
- Claims submission and processing, automated follow-ups, collections and remittance advice
- Managing the schedules of employees of the clinic
- Managing the inventory of medical supplies and office materials
- Inter and Intra clinic/location communication
- HIPAA compliance
- Patient portal

In a Cloud-based Practice Management solution, the software is accessed through a browser over the Internet, which is why continuous Internet availability is crucial. Cloud-based Practice Management solutions can be customized, although admittedly to a limited extent when compared to on-site solutions. However, Cloud-based solutions can effectively compete with the on-site solutions regarding functionality and features.

A patient portal is essentially a website on which patients can log in and access most portions of their medical records. In a patient portal, the information is presented in a manner which makes it easy for the patient to understand the nature of his conditions and his treatment. Most Patient Portals also provide some level of interactivity with the healthcare provider.

Patient Portals represent one of the key advantages of Cloud-based systems over in-house systems. A Patient Portal interfaces with sensitive, private information of individuals, which should be securely available and web-accessible at all times. It is, therefore, best if the responsibility of managing the patient portal is assigned to an entity that has the expertize, the infrastructure, the resources, and experience to develop and achieve high availability, high-security information systems. Most clinics or physician

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

practices do not have this capability, nor is it practically feasible to invest in infrastructure and human resources to undertake such capability. It is, therefore, best to opt for a Patient Portal that is Cloud-based.

Many Practice Management systems and EMR systems include ePrescription capability as a part of their overall functionality. An ePrescription system is a computerized system in which the prescription is either entered by the physician/nurse practitioner or generated by information available to the system. The prescription can be automatically communicated to pharmacies associated with the healthcare provider. Further, ePrescription systems also have inbuilt rules/databases about drug-drug allergies and appropriateness of each drug in the context of various health conditions.

An ePrescription system brings several benefits (Soman, 2011):

- It eliminates the possibility of a prescription being incorrectly interpreted due to poor handwriting, thereby reducing medical errors.
- It reduces the occurrence of drug-drug allergies and the possibility of the inappropriate drug being administered in the view of the patients' comprehensive medical history.
- It makes ordering refills easier since prior prescriptions can be easily accessed electronically.
- ePrescriptions are also found to increase compliance by patients merely because they are more convenient

A Cloud-based ePrescription system is likely to be much more highly available and secure and thus be able to communicate prescriptions to pharmacies with a higher reliability than an in-house system. Secondly, a Cloud-based system is more likely to cooperate with more and more pharmacy chains over time, enhancing the number of pharmacies the ePrescriptions can be sent to.

Even if a clinic does not use Cloud-based systems for practice management, it is a brilliant idea to consider Cloud-based systems for data backup and secondary storage.

The way most Cloud-based backup systems work is as follows (Soman, 2011).

There is software installed on one or more of the computers within the facility (called the backup client) connected to the primary storage. The backup client tracks the changes to the data stored on the primary storage at regular intervals and communicates the changes to the Cloud-based storage space of the service provider.

The backup client encrypts all information before it is transmitted to Cloud storage. Secondly, bandwidth requirements are significantly reduced because the client only communicates the changes that have occurred in the data since the last backup run, rather than transmitting all the details again.

A notable special case of a data backup system, due to the large sizes of medical images, is the Picture Archival and Communication System ('PACS'). In the process of administering health care, various images, graphs, pictures, and scans are generated. That includes images from X-Rays, ultrasound machines, MRI and CT scans, endoscopy, mammograms, computed radiography, etc. Even scanned images of paper documents are often required to be stored when legacy paper-based data is digitized via scanning. To ensure uniformity, a standard format used to store and communicate pictorial information has been defined—the Digital Imaging and Communications in Medicine ('DICOM').

Cloud-based laboratory solutions are available for various functions such as microbiology, histology and cytology processes, Synoptic/Antibiogram reporting, Image/Document capture, Inventory Management, Voice recording, Blood bank. They have all the functionality of on-site solutions.

Cloud-based Radiology systems are similar to PACS systems. These systems support basic Radiology workflows where physicians can register examination requests for existing or new patients, and select ap-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

propriate billing rules. It is also possible to interface laboratory devices directly to Cloud-based systems. That means that the readings from instruments or captured images are directly uploaded to the cloud.

Overall, cloud-based healthcare IT systems are easier to get started; distinctly cost-effective when compared to on-site solutions; more convenient, especially for smaller clinics and private physician practices, while allowing doctors to collaborate on any medical data easily, and work on the system from anywhere, including from home.

Cloud computing is a novelty that rapidly showed tremendous opportunities for application in medicine and health care improvement (Archondakis, Pouliakis, Margari, & Karakitsos, 2014; Eugster, Schmid, Binder, & Schmidberger, 2013; Glaser, 2011; Kuo, 2011; Lupse, Vida, & Stoicu-Tivadar, 2012; Mirza & El-Masri, 2013; Moumtzoglou & Kastania, 2014; Patel, 2012; A. Pouliakis, Archondakis, Karakitsou, & Karakitsos, 2014; A. Pouliakis, Spathis, et al., 2014; Rosenthal et al., 2010; Waxer, Ninan, Ma, & Dominguez, 2013; Webb, 2012). It is forecasted that there will be an increase in the US cloud computing market for medical images approximately 27% by 2018 at a Compounded Annual Growth Rate (CAGR). That is mainly due to the growing volume of medical images and the increasing costs of the ownership for owning Picture Archiving and Communication Systems (PACS) (GlobalData, 2012).

Big Data

Big Data is a terminology which is used to describe a collection of datasets being so broad and complicated that it becomes challenging to process them using the standard and on-hand database management tools or the available traditional data processing software applications. Although there is no single formal definition of Big Data, most definitions seem to associate to four Vs which include volume, velocity, variety, and veracity. Volume relates to the unusually large data which are produced, collected and exchanged; velocity refers to the need for real-time analysis; variety identifies the different forms of data that are difficult to integrate while veracity emphasizes the issues which are related to trust and uncertainty. However, not all data that occupy a large volume are considered big. Thus, when defining Big Data, the size is not the primary characteristic; the most appropriate characteristic is the number of independent data sources (separate databases), each one having the potential to interact and be interconnected with one or more data sources.

According to the newest market studies (Market Research Future, 2017; Markets and Markets, 2017), the revenue of Big Data by 2020 and 2021 is estimated in the range of 25-35 billion USD. Thus, the Big Data market is continuously increasing. It is a very competitive market, and many stakeholders have presented innovative products and services dealing with Big Data. The Big Data ecosystem has several discrete areas: a) storage and management including data servers, software, networking, monitoring, administration and data management b) tools and technologies specialized for Big Data analytics that generates insight from collected data c) healthcare intelligence for the organization and d) embedding the Big Data framework into the organization's business intelligence infrastructure (Koumpouros, 2015).

There are several Big Data solutions either open source or commercial ones. Open platforms have advantages as they come at no cost for software acquisition and permit the use of various tools for data analytics. However, they require significant effort to install and setup by specialists. The Apache Hadoop framework is the pioneering and mostly used open platform. Other open source systems include NoSQL databases: MongoDB and Cassandra, among others. The Hadoop framework with proper scaling can support the analysis of petabytes of data. In this arena, another platform, Red Hat Enterprise Linux, in companion with the OpenStack hybrid cloud, supports Big Data applications. In the commercial prod-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

uct arena, Info Sphere is one of the products offered by IBM, another platform based on the Hadoop system. It includes data warehouse and stream computing software to perform real-time analysis on huge volumes of streamed data. Cisco offers the Cisco Common Big Data Platform, which is based on the Cisco's Unified Computing Architecture and concept. Due to the company's networking history, it integrates computing, networking and storage capabilities. Other Big Data platform and services vendors include Cloudera, Hortonworks, Platfora, DataStax, 10Gen, and Amazon, being supporters of Hadoop and NoSQL. As well as Hadapt, Platfora, Splunk, Datameer, and Karmasphere.

One of the major Big Data and cloud service providers are Amazon.The main service is the Elastic MapReduce (EMR), a Hadoop based platform that scales to thousands of nodes per cluster. There is support for easy to use mission –critical, real - time applications, a petabyte scale warehouse service. Finally, it includes business intelligence tools and can operate on datasets of a few hundred gigabytes to petabytes.

Healthcare is probably one of the leading areas in the Big Data arena, because significant amounts of personal and sensitive, health-related data can be continuously collected from patients. In the medical field, those data are collected by different actors, in different locations, for various purposes and is used/viewed differently by each information consumer (Koumpouros, 2015). For the medical specialists, to reach the best possible diagnosis and subsequently, the appropriate therapeutic scenario, accurate and complete information is required; without Big Data, this is not feasible.

Evidence-based medicine (EBM), is another factor that stresses the need for Big Data in healthcare, as it relies on large scale data sets, usually distributed (Beckmann & Lew, 2016; Sniderman, LaChapelle, Rachon, & Furberg, 2013). EBM is used by the physicians as well as by health care management and insurance organizations, paving the way for the acceptance of Big Data.

As far as population health management is concerned, there is already an application of Big Data (Wells et al., 2016). Specifically, the application is mainly based on the U.S. Medicare and Medigap insurance programs. Within this application, data from diverging lines of evidence are used:

- Demographic and contact details along with insurance coverage.
- Insurance claims, submitted by doctors, hospitals, and other providers to Medicare; those data include standardized codes for medical diagnoses and procedures, place of treatment, amounts billed, as well as whether service was provided:
    - In the emergency room
    - Inpatient service
    - Ambulatory service iv)
    - Laboratory
    - Long-term care unit
    - Ancillary service
- Prescribed pharmaceuticals that include the drug name and class, a drug code identifier (applied at the national level), drug dosage, dates, and the number of days for drug usage. It is worth noting that this is a separate database linkable to the medical claims and administrative data from Medicare.
- A questionnaire is measuring experiences with health plans of Medicare which includes the health status, satisfaction, and experiences from healthcare services.
- U.S. Census data which are related to the zip code and are used to identify differences by in the geographic regions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Data from Dartmouth Atlas which describes the supply of healthcare services in a geographically geographical context.
- The AmeriLINK Data Sourcing system which indicates the social and economic status of members.

The processed Medicare data were around 4 million records, and the results of health analytics proved to be very interesting. The aim was to 'get the right person to the right care at the right time.' Individuals that filled the questionnaire had two times more probability to use a Medicare service while increased time in care coordination was linked with fewer hospital readmissions. Thus, those individuals were more likely to visit a physician's office than return to the hospital and increase healthcare costs. Improved depression management enabled reduction of depression symptoms in about 60% of engaged individuals in the program. Finally, Big Data analytics enabled suggestions for participation in a wellness program.

However, there are still numerous issues in the Big Data arena, especially in healthcare. By 2020, experts predict that data will be produced at a rate of about 40 trillion gigabytes per day. Current information technology is not capable of storing and processing such large volumes. Also, there is a shortage of skilled personnel to collect, manage and process these data. Furthermore, there is resistance from health providers and administrators, as they do not understand the benefits of Big Data and especially for population health management. For data collected from the medical facilities, systems' time and technical diversity seem to be a barrier for data collection. Moreover, medical systems are constantly changing, as they adapt to new technological and scientific advances, as well as to emerging organizational and regulatory requirements and newly developed standards. As a sequel, even in a single hospital there are several information systems and often a bridge is required to ensure interoperability. Despite the evolution in medical standards, it seems that several efforts are still needed. Within the healthcare facilities, the replacement of existing systems depends on the willingness and financial capacity of healthcare providers. Additional barriers will emerge due to the application of M-Health; whereas data collection will be more intense, individuals are already able to collect them from their own devices.

M-Health

M-Health describes the use of a broad range of telecommunication and multimedia technologies within the wireless care delivery design and can be broadly defined as the provision of healthcare services via mobile communication devices (Dimitriou et al., 2017; Metelmann & Metelmann, 2017; Rana, 2017). M-Health establishes healthcare communities in which every stakeholder can participate (O'Connor, Heavin, & O'Donoghue, 2015). However, it disrupts the traditional service model where healthcare information, security, and access is centrally managed, maintained and limited, transforming the health sector and destroying components that are slow to adapt.

M-Health interventions range from simple to complex applications and systems that remotely coordinate and actively manage patient care (Abraham Pouliakis, Archondakis, Margari, & Karakitsos, 2016). In this context, it offers an elegant solution to the problem of accessing the right information where and when it is needed within highly fluid, distributed organizations. Moreover, it removes geography and time as barriers to care by establishing connectivity with remote locations and remote workers, creates new points of contact with patients, and changes the frequency and intensity of health care delivery. It also establishes effective new treatment modalities like telehealth, remote patient monitoring, self-care, and home health while it blurs the boundaries between professional medical advice and self-care. Overall, M-Health blends three bodies of knowledge: high technology, life sciences, and human factors.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

M-Health could also be thought as the practice of medicine and public health supported by mobile devices. The use of mobile computing and communication devices, such as mobile phones, tablet computers and personal digital assistants (PDAs), by health professionals, is nowadays rapidly growing. Wireless networks are the answer for the interconnection of ubiquitous devices. However, wireless communications increase security challenges, privacy concerns, and trust issues for the wireless network, the mobile devices, the connected medical devices and e-health applications (Grami & Schell, 2004). Wireless, mobile nodes is a new challenge in the distributed systems which carry out functions like packet forwarding, routing, and network management (Chlamtac, Conti, & Liu, 2003). Various protocols and algorithms are available for mobile ad-hoc networks (Chlamtac et al., 2003) while the quality of service models for mobile ad-hoc networks and network security are significant challenges to be surveyed (Chlamtac et al., 2003). As a result, different quality of service constraints for various services illustrates the need for an automatic wireless technology option (Janevski, 2009; Kitanov, Monteiro, & Janevski, 2016).

Currently, the generations of mobile phone networks (3G/UMTS, 4G and the predecessor of 5G, i.e., 4.5G) are packet-based architectures (Kastania & Moumtzoglou, 2012). Mobile telephony since the 3G technology enabled the seamless delivery of web content to mobile consumers (Wilson, 2006). Users' management involves subscriptions and profiles for different terminals (cell phone, PDA, laptops, desktops) (Hong & Leon-Garcia, 2005) while the choice of a cell/mobile device is heavily influenced by the provided mobile services and device capabilities.

Existing cellular and Wi-Fi networks that have, nowadays, merged with the Internet, represent the 4th Generation (4G) and the predecessor of 5G (i.e., 4.5G) wireless, mobile Internet networks (Gani, Li, Yang, & Zakaria, 2008). 4G generation of wireless enhanced and replaced the 3G systems. 4G is a mobile network, IP-based, providing a connection via always the best available network using seamless roaming and free radio access technologies. In 4G mobile systems, different access technologies are combined, in the best possible way, for various radio environments and service requirements. 4G technologies have already implemented much larger data rates supporting full mobility while enabling wireless connection and access to multimedia services with high-quality voice and high-definition video. 4G wireless telecommunication supports smart multi-media providing higher bit rate and wider bandwidth, 3D image technology, streamed HD TV with remote control, video multimedia and advanced multimedia messaging service. Mobile multimedia provision requires a modern mobile communication system to make ubiquitous communication using mobile characteristics. High-quality voice, video, and data services can be delivered on a metropolitan using Wi-Fi (IEEE 802.11a/b/g/n) as a transition towards 5G. Mobile WiMAX, while expected to fulfill the goals of 4G technology offering full broadband for all services, is not, nowadays, the technology of choice. In this context, global roaming and open interconnection interfaces are essential functions of 4G communications systems.

According to many researchers, mobile applications ensure seamless mobile services for end-users, anywhere, and on any device. However, the shifts towards more advanced networking systems depend on the development of wireless radio access technologies and networking technologies.

Services adding to 4G include 5G, which is expected to be a more intelligent technology for the whole interconnection world without limits. 5G technology is currently implemented and is the standard of mobile telecom providers; it is synonym with the LTE Advanced (Long Term Evolution) and MIMO (multiple input / multiple outputs) technologies that enable the multiplication of capacity and frequency use along with optimization in the urban and indoor environments to handle multi path effects.

5G is planned to provide additional services and have more advantages, related to energy consumption, network coverage, spectrum handling, bandwidth use, and allows users to participate in a fully distributed

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

or cellular-controlled fashion. All IP-based services for heterogeneous wireless access technologies, assisted by mobile IP, are to be provided by 5G mobile communication networks. 5G aims at a higher capacity and simultaneously a higher density of mobile users. Moreover, it supports device-to-device, and machine to machine reliable communications (i.e., mobile ad-hoc networks / MANETs). Thus, 5G aims at lower latency than the 4G equipment and lower battery consumption and is an enabling technology for the Internet of things. Nowadays, there is no standard for the deployment of 5G. However, the expected characteristics include:

- Tens of megabits per second, simultaneously for dozens of thousands of users.
- More than 100 megabits per second for metropolitan networks.
- 1 Gb per second, simultaneously available to all users on the same building floor.
- Capacity for hundreds of thousands of simultaneous connections for the implementation of Internet of Things.
- Spectral efficiency / i.e., more Mbps for the same frequency slot
- Improved coverage
- Significantly reduced latency

MEDICAL APPLICATIONS FOR THE NEXT GENERATION OF MOBILE DEVICES AND NETWORKS

Mobile web services represent the integration of web services between mobile devices without any centralized entities (Srirama, Jarke, Zhu, & Prinz, 2008). Factors influencing mobile hosts are network characteristics, searching capability of the services and computational resources (Srirama, 2006) while collaborative sharing of peer resources is possible for Mobile Hosts in P2P (Srirama, 2006). Mobile Host in JXME JXME allows collaboration, image sharing, and location-based services (Srirama, 2006). Each Mobile Host can provide some services in the wireless network acting, through the cell phone, as both web service client and provider (Srirama, Jarke, & Prinz, 2007). The quality of service (QoS) for the Mobile Host is critical regarding security and scalability.

A mobile phone is a communications device offering wide-area networking capabilities (Roussos, Marsh, & Maglavera, 2005). Cell phones can serve as control devices for ubiquitous systems and information service endpoints with a focus on protecting the user's information (Roussos et al., 2005). Many mobile phones interact with other nearby devices due to the availability of wireless networking capabilities (Itani, Kayssi, & Chehab, 2016; Roussos et al., 2005). Visual tag clicking on camera phone allows user interaction with a mobile service (Toye et al., 2007) while mobile content can be downloaded, in commercially available products scanning visual tags (Toye et al., 2007). Moreover, data rates increase web browsing and offer reduced access times. However, there are constraints associated with small mobile displays and keypads (Toye et al., 2007) while larger data rates increase energy consumption and complexity of the device.

Overall, the smart Internet phone lets the mobile phone select the best connections (Janevski, 2009). Furthermore, all smartphones (cell phones that include PDA functionality) are equipped with the functionality to access high-speed data through a variety of networks including those supported by mobile phone networks (3G and 4G+), local networks (WiFi) and personal area networks (Bluetooth and NFC). Applications for the next-generation of cell phones will provide localized, real-time services and local-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ized information services. However, confirmation is needed of the services provided by the smartphones in a wireless network (Srirama et al., 2007).

Handheld computing has already numerous applications such as:

- Ambulatory medicine (K. Banitsas, Perakis, Koutsouris, Konis, & Tachakra, 2005; K. A. Banitsas, Perakis, Tachakra, & Koutsouris, 2006; Kiselev, Gridnev, Shvartz, Posnenkova, & Dovgalevsky, 2012; Pavlopoulos, Kyriacou, Berler, Dembeyiotis, & Koutsouris, 1998; Rosales Saurer, Mueller-Gorchs, & Kunze, 2009; Zerth, Besser, & Reichert, 2012).
- Diabetes management (Quinn et al., 2011; Ribu et al., 2013; Skrovseth, Arsand, Godtliebsen, & Hartvigsen, 2012; Spat et al., 2013).
- Asthma management (Finkelstein, Hripcsak, & Cabrera, 1998; Gupta, Chang, Anyigbo, & Sabharwal, 2011).
- Control of obesity (Patrick et al., 2009).
- Smoking control (Ghorai, Akter, Khatun, & Ray, 2014; Ybarra, Holtrop, Prescott, & Strong, 2014).
- Seizure management (Pandher & Bhullar, 2014).
- Stress management (Clarke et al., 2014) and
- Treatment of depression (Burns et al., 2011) among others.

In the field of M-Health most applications are fitness related (43%) followed by health resource (15.0%) and diet/caloric intake (14.3%,), while the user engagement has the form of self-monitoring and training (74.8%) (Sama, Eapen, Weinfurt, Shah, & Schulman, 2014). On the other hand, despite the applications which target patients, currently, there are rather limited applications targeting physicians and doctor-patient interactions (Martin, 2012). The pioneering field seems to be radiology consultation for X-rays and mostly Computer Tomography (Choudhri et al., 2012; Choudhri et al., 2013; Johnson et al., 2012; Toomey et al., 2010) and ECG transmission (Vaisanen, Makijarvi, & Silfvast, 2003).

However, the mobile device has emerged recently as a technology gadget collecting biomedical data, either in the form of the mobile phone or as a specialized appliance, usually in the shape of a wrist attached device. This plethora of devices has created a new movement, called quantified self (QS) (Fawcett, 2015). QC involves ordinary and usually healthy people that record and process aspects of their lives with the help of (usually) self-collected data. This QS movement has attracted the interest of commercial companies offering products and services (hardware and software). The M-Health applications being of interest to physicians, patients, and the science of individuality include the following.

Mobile MIM (MIM Software, 2017) is a software application for viewing MRI, PET and CT data using mobile devices compatible with the Apples' iPhone series operating system (iPad®, iPhone®, and iPod touch®). MRI, PET, and CT images are transmitted using software, and it is possible to measure the image intensity and distances.

MobiUS™ is an image based series of products (Mobisante Inc., 2017) supporting ultrasound (US) imaging. That is smartphone and tablet-based and is cheaper than traditional US systems; being equipped with specialized monitors. Moreover, this system supports sharing of images among patients and clinicians.

Gyromaniac (Subversus Interactive, 2011) uses the iPhone gyroscope electronic components to help physicians practice spatial orientation: for example, to assist gastroenterologists during training how to perform a colonoscopy (Iltifat, 2010).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

There are, as well, software platforms for continuous wearable, non-obtrusive ECG data (cardiac) monitoring. Such solutions can be of help for cardiologists to monitor and manage patients with angina, infarction of the myocardium, cardiac interventions such as stents, pacemakers, and bypass, as well as cardiac failures and ambulatory medicine. Such devices have been in service for gynecologists for fetus monitoring (Medical design, 2013; Su & Chu, 2014). That avoids frequent visits to the outpatient department for follow up, the unnecessary waiting time and anxiety for patients. That system also aids proactive management of complications, identifying them with intelligent analysis as and when they occur and addressing them based on evidence.

A recently announced device (Wavelet Health, 2017) incorporates a photoplethysmograph (PPG) sensor; a pulse oximeter illuminating the skin and measuring changes in light absorption; such a device is capable to continuously monitor perfusion of blood to the dermis as well as subcutaneous tissue; therefore, it can detect heart rate as well as blood oxygen levels.

Recent announcements (University of Illinois College of Engineering, 2017) are related to a lab on a chip device that when coupled with a smartphone can perform lab-grade medical diagnostics. Such tests usually require large, specialized and expensive instruments. However, the developed inexpensive microchip incorporates technology common in the three top categories of medical diagnostics. The microchip promises to perform thousands of already available tests. It is based on the mobile phone camera that is used as a spectrometer. Specifically, the biological sample should be in a fluid form; this is illuminated by the cell phone's white LED flash. Subsequently, the light from the fluidic sample is collected by an optical fiber and after being captured by the phone's internal camera can be evaluated by the use of the software. Such technology enables the application of advanced biological tests in remote areas, often in the developing world, and personalized tests at the individual patient level.

In the field of otolaryngology some applications can determine reference sound levels with a minimum l variation; therefore, it is feasible to conduct a hearing screening on devices calibrated with a predefined reference noise level (Masalski, Kipiński, Grysiński, & Kręcicki, 2016).

Also, the advanced communication capabilities allow voice and video conferencing features for patient/doctor and doctor/doctor interactions as well as seamless transmission of captured images and video.

Mobile electroencephalograph (EEG) is another area that mobile devices may become a commodity. Emotiv Insight Brainware (see Figure 1) is a five channel EEG device that can capture brain signals and transmit them via Bluetooth interface to the mobile phone (EMOTIV, 2017). Such commercial and portable EEG devices have already been used to measure auditory potential on children and adults (Barham et al., 2017). They are used to convert brain signal to computer commands, and some identify the mental status and record it for archiving purposes.

Furthermore, modern high-end mobile phones are coupled with sensors that can measure life signals and communicate with commercial devices that measure health signals such as blood glucose. For instance, some cell phones have integrated heart rate sensors with high accuracy and step counters. They also have applications (Figures 2-4) that measure and store data to the cloud, such as exercise and sports activity, heart rate, stress, sleeping duration, consumption of food and liquids, weights and glucose when coupled with appropriate medical devices or when the user wishes to supply such information.

Conclusively, the recent advances of the mobile telephony have contributed to the conversion of the mobile telephones into smartphones; due to the sophisticated operating systems. Especially, during the previous five years, hardware components enriching cell phones became available, including more processing power (8 or more processors), high-definition cameras (more than 15Mpixel) and large display resolutions. Moreover, the 4.5G and WiFi connectivity have created an always open, rather inexpensive,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A commercial portable EEG device capable to communicate via Bluetooth with the mobile phone

communication channel with the Internet. Today, there is no more the trend to shrink the telephone; in contrast, mobile phones are becoming larger and complete tablets regarding size and weight; they are easy to use and have capabilities similar to computers available five years ago. It seems that the mobile devices have now the maturity of computers that were used in the health sector five years ago.

Issues, Controversies, Problems

First, we must recognize that to center population health efforts on the patients we are overlooking those who are not patients. Even for those who are exceptionally engaged patients, their interaction with the healthcare system is a mere fraction of health and well-being influences. Consequently, if we design population health efforts solely around patients, we fail to address the enormously significant social determinants of health which include environmental, economic, and cultural factors.

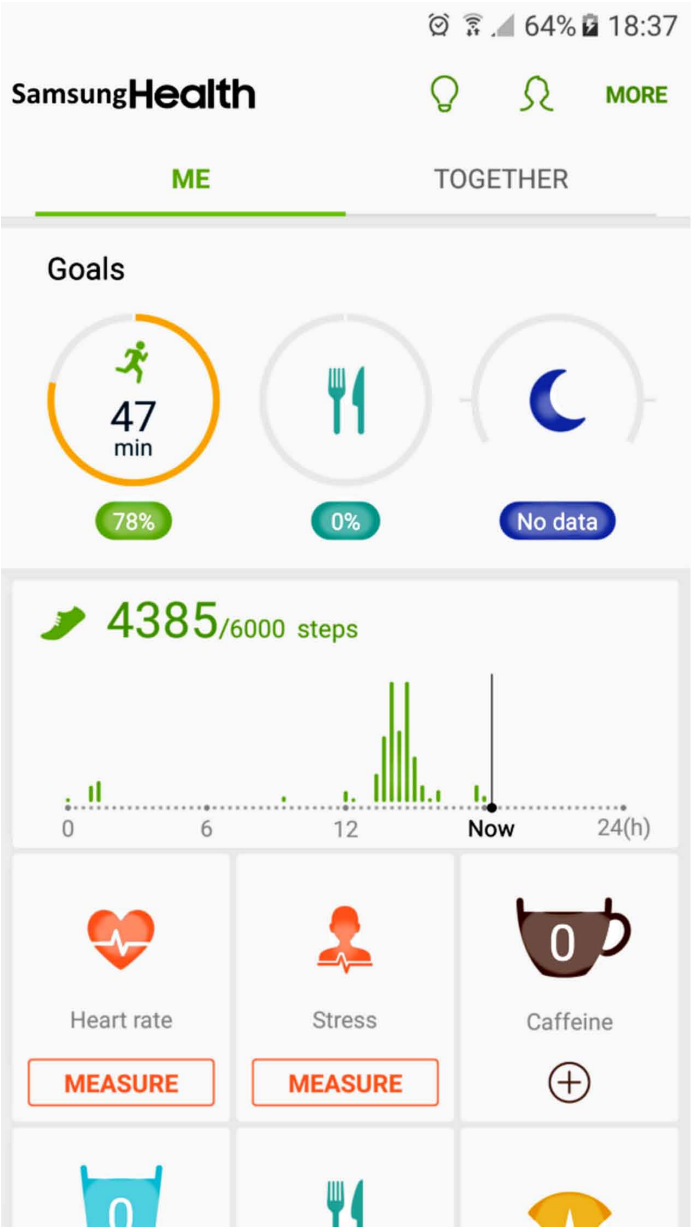
Instead, we need a proactive, holistic, longitudinal, without walls, deeply personal approach to population health management. That means the approach should include not only the care of sick but also well care, as the goal should be lifelong health and quality of life support. That also includes proactively reaching out to hard-to-reach populations. Moreover, the approach should:

- Address the range of psychological, biological and social factors that influence health and wellness.
- Leverage the patient as a source of continuity by engaging him or her as an integral member of the care team.
- Recognize that supporting an individual's lifelong health and well-being cannot be relegated to anyone particular setting.
- Take into account the personal goals, strengths and challenges as for primary drivers of care management, planning and decision-making.

While 'Big Data' has its place for identifying patterns and gaps in care and predictive analysis, the most relevant data regarding the science of individuality are self-reported data. In this context, the population health management strategies will fail to engage patients, reduce costs, and improve outcomes, if they are implemented without the underpinning of person-centered care. Ultimately, the health of a population has nothing to do with data and technology but with the individuals that make up the population.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

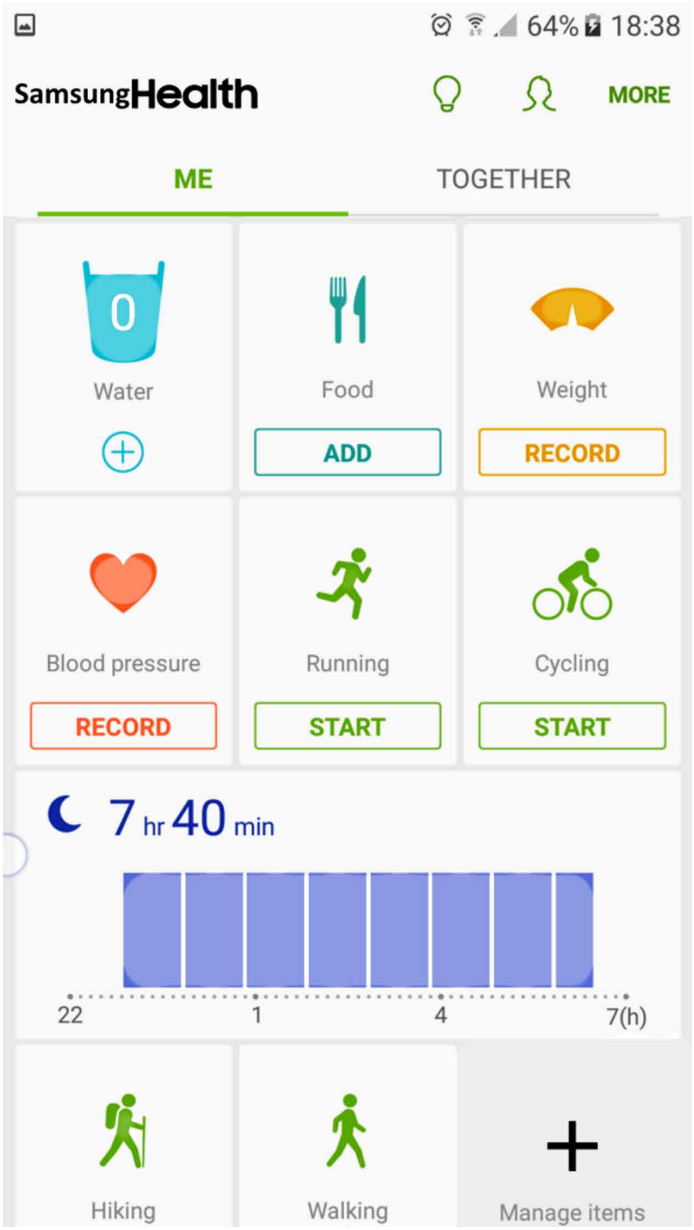
Figure 2. A display of mobile phone applications for the measurement of health and well-being data (part 1)



To this direction, and due to the sensitivity of the stored and transferred information, data security for all of the associated technologies is of essential importance (Spyra, Buchanan, Cruickshank, & Ekonomou, 2014). In this context, the European Union approved (in April 2016) the General Data Protection Regulation (GDPR); the most significant change in the European legal framework the last twenty (20) years. The GDPR will have a direct application in all member states on May 25, 2018, and enforcement is assured by heavy fines on non-compliant organizations. However, the implementation of the GDPR,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

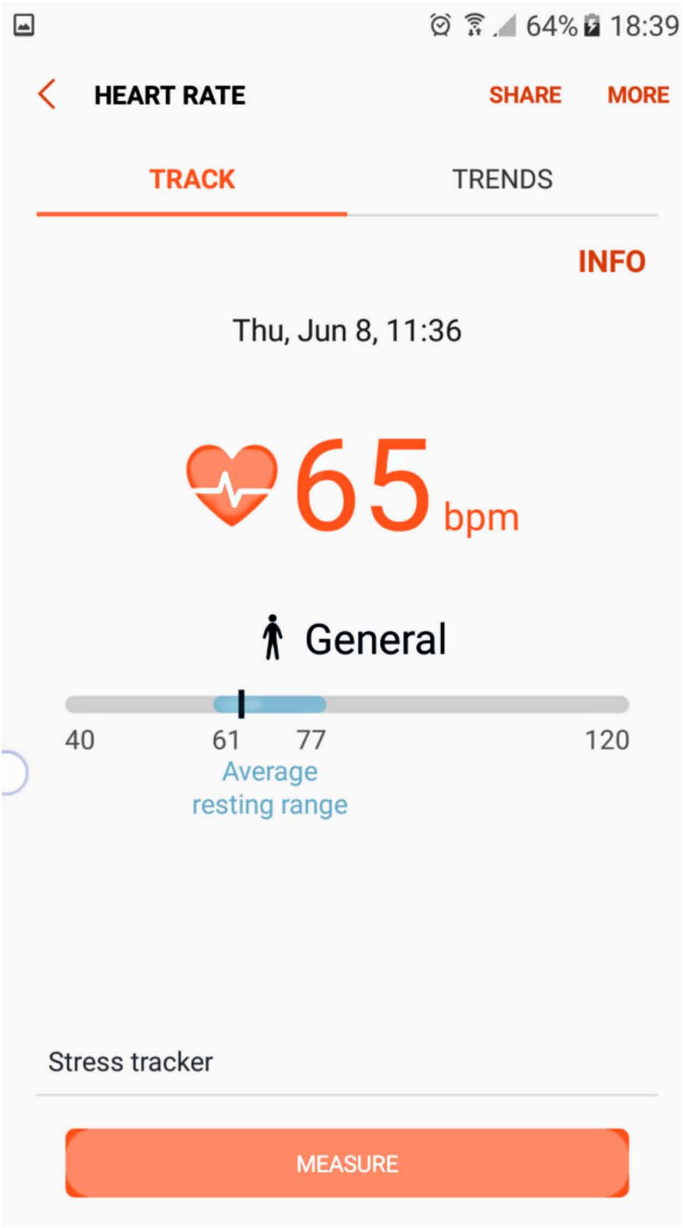
Figure 3. A display of mobile phone applications for the measurement of health and well-being data (part 2)



to a large scale, might be a big issue, as information flowing across the EU borders and data storage in non-EU countries could be essential, taking into account that non-EU countries are not obliged to apply the GDPR. As a result, services offered by organizations residing outside the EU could be vital for e-health applications targeting individuals.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. A display of mobile phone applications for the measurement of health and well-being data (part 3)



FUTURE RESEARCH DIRECTIONS

The future of population health management is tied to the rise of cognitive computing. Cognitive computing is the simulation of human thought processes in a computerized model and uses parallel processing and artificial intelligence to convert unstructured into structured data, search the medical literature, and find connections among multiple types of evidence, in short, it creates knowledge.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The goal of cognitive computing is to build automated IT systems that are capable of solving problems without requiring human assistance. The cognitive computing systems:

- Use natural language processing and machine learning algorithms.
- Mine the data for information.
- Refine the way they search for patterns to become capable of anticipating new problems and modeling solutions.

Heretofore, cognitive computing is used in the following artificial intelligence (AI) applications:

- Expert systems
- Natural language programming
- Neural networks
- Robotics
- Virtual and augmented reality

CONCLUSION

Population health management (PHM) has been, so far, a discipline which studies and facilitates care delivery across a group of individuals or the general population. In the context of population health management, the life science industry has no motivation to design drugs or devices that are only effective for a distinct segment of the population. At the same time, the regulatory agencies are entirely risk-averse and are suppressing remarkable innovative and even frugal opportunities to change medicine.

However, the major outgrowth of the science of individuality, as well as the rising 'wiki medicine', that is, the individualized medicine which respects individuality, fully recognize the uniqueness of the individual. In this regard, it will not be long until digitizing a person unlocks the root cause for what is wrong, creating valuable knowledge that can save or markedly improve the quality of life.

Cloud computing, Big Data and M-Health technologies offer the resources to succeed in dealing with the shortcomings of the population health management approach. Specifically, while population health management focuses on what happens on the bed or a doctor's office., the technologies above facilitate the propagation of the science of individuality by taking the challenges of wellness, prevention and lower cost. Finally, cognitive computing is considered the future of population health management.

REFERENCES

Archondakis, S., Pouliakis, A., Margari, N., & Karakitsos, P. (2014). The Implementation of Cloud-Based Telemedical Applications for External Quality Control Purposes in the Field of Cytopathology. *International Journal of Reliable and Quality E-Healthcare*, *3*(3), 22–36. doi:10.4018/ijrqeh.2014070102

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Banitsas, K., Perakis, K., Koutsouris, D., Konis, G., & Tachakra, S. (2005). Ambulance 3G. *Conference Proceedings; ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* (Vol. *1*, pp. 371–374). doi:10.1109/IEMBS.2005.1616422 PMID:17282191

Banitsas, K. A., Perakis, K., Tachakra, S., & Koutsouris, D. (2006). Use of 3G mobile phone links for teleconsultation between a moving ambulance and a hospital base station. *Journal of Telemedicine and Telecare*, *12*(1), 23–26. doi:10.1258/135763306775321335 PMID:16438775

Barham, M. P., Clark, G. M., Hayden, M. J., Enticott, P. G., Conduit, R., & Lum, J. A. G. (2017). Acquiring research-grade ERPs on a shoestring budget: A comparison of a modified Emotiv and commercial SynAmps EEG system. *Psychophysiology*, *54*(9), 1393–1404. doi:10.1111/psyp.12888 PMID:28497557

Beckmann, J. S., & Lew, D. (2016). Reconciling evidence-based medicine and precision medicine in the era of big data: Challenges and opportunities. *Genome Medicine*, *8*(1), 134. doi:10.118613073-016-0388-7 PMID:27993174

Care Continuum Alliance. (2012). Advancing the Population Health Improvement Model. Retrieved 2014 from http://www.fiercehealthit.com/story/hennepin-health-project-looks-build-countywide-ehr-program-national-implica/2012-01-10

Chlamtac, I., Conti, M., & Liu, J. J. N. (2003). Mobile ad hoc networking: Imperatives and challenges. *Ad Hoc Networks*, *1*(1), 13–64. doi:10.1016/S1570-8705(03)00013-1

Choudhri, A. F., Carr, T. M. III, Ho, C. P., Stone, J. R., Gay, S. B., & Lambert, D. L. (2012). Handheld device review of abdominal CT for the evaluation of acute appendicitis. *Journal of Digital Imaging*, *25*(4), 492–496. doi:10.100710278-011-9431-9 PMID:22146833

Choudhri, A. F., Norton, P. T., Carr, T. M. III, Stone, J. R., Hagspiel, K. D., & Dake, M. D. (2013). Diagnosis and treatment planning of acute aortic emergencies using a handheld DICOM viewer. *Emergency Radiology*, *20*(4), 267–272. doi:10.100710140-013-1118-8 PMID:23525945

Clarke, J., Proudfoot, J., Birch, M. R., Whitton, A. E., Parker, G., Manicavasagar, V., ... Hadzi-Pavlovic, D. (2014). Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: Secondary analysis of a randomised controlled trial. *BMC Psychiatry*, *14*(1), 272. doi:10.118612888-014-0272-1 PMID:25252853

Coughlin, J. F., Pope, J. E., & Leedle, B. R. Jr. (2006). Old Age, New Technology, and Future Innovations in Disease Management and Home Health Care. *Home Health Care Management & Practice*, *18*(3), 196–207. doi:10.1177/1084822305281955

Cusack, C. M., Knudson, A. D., Kronstadt, J. L., Singer, R. F., & Brown, A. L. (2010). *Practice-Based Population Health: Information Technology to Support Transformation to Proactive Primary Care* (p. 44). Retrieved from https://pcmh.ahrq.gov/sites/default/files/attachments/Information%20Technology%20to%20Support%20Transformation%20to%20Proactive%20Primary%20Care.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Dimitriou, M., Giazitzi, K., Stavridi, E., Palisidis, G., Karathanos, V., & Boskou, G. (2017). Evaluation of an Electronic Application for Nutritional Information in Food Service Outlets: A Pilot mHealth Application. *International Journal of Reliable and Quality E-Healthcare*, *6*(1), 46–58. doi:10.4018/IJRQEH.2017010104

EMOTIV. (2017). Emotiv insight Brainware. Retrieved Sept. 2, 2017, from https://www.emotiv.com/insight/

Eugster, M. J., Schmid, M., Binder, H., & Schmidberger, M. (2013). Grid and cloud computing methods in biomedical research. *Methods of Information in Medicine*, *52*(1), 62–64. PMID:23318694

Fawcett, T. (2015). Mining the Quantified Self: Personal Knowledge Discovery as a Challenge for Data Science. *Big Data*, *3*(4), 249–266. doi:10.1089/big.2015.0049 PMID:27441406

Finkelstein, J., Hripcsak, G., & Cabrera, M. (1998). Telematic system for monitoring of asthma severity in patients' homes. *Studies in Health Technology and Informatics*, *52*(Pt 1), 272–276. PMID:10384460

Gani, A., Li, X., Yang, L., & Zakaria, O. (2008). Radio frequency combination for TCP/IP suite protocol improvement in 4G mobile internet networks. *International Journal of Communication*, *2*(1), 27–36.

Ghorai, K., Akter, S., Khatun, F., & Ray, P. (2014). mHealth for Smoking Cessation Programs: A Systematic Review. *Journal of Personalized Medicine*, *4*(3), 412–423. doi:10.3390/jpm4030412 PMID:25563359

Glaser, J. (2011). Cloud computing can simplify HIT infrastructure management. *Healthcare Financial Management*, *65*(8), 52–55. PMID:21866720

GlobalData. (2012). Snapshot: The US cloud computing market for medical imaging. Retrieved 22/08/2013, from http://www.medicaldevice-network.com/features/featuresnapshot-the-us-cloud-computing-market-for-medical-imaging

Grami, A., & Schell, B. H. (2004). Future trends in mobile commerce: Service offerings, technological advances and security challenges. *Paper presented at the Second Annual Conference on Privacy, Security and Trust*.

Gupta, S., Chang, P., Anyigbo, N., & Sabharwal, A. (2011). mobileSpiro: accurate mobile spirometry for self-management of asthma. *Paper presented at the First ACM Workshop on Mobile Systems, Applications, and Services for Healthcare*, Seattle, Washington. 10.1145/2064942.2064944

Hillman, M. (2002). Testimony before the Subcommittee on Health of the House Committee on Ways and Means, Hearing on Promoting Disease Management in Medicare. Retrieved 10 June, 2015, from http://waysandmeans.house.gov/Legacy/health/107cong/4-16-02/4-16hill.htm

Hong, J. W., & Leon-Garcia, A. (2005). *Paper presented at the 19th International Telegraffic Congress*.

Howe, R., & Spence, C. (2004). *Population Health Management: Healthways" PopWorks* (Vol. 2, pp. 291–297). HCT Project; Retrieved from http://mthink.com/article/population-health-management-healthways-popworks/

Hsu, C. H., Lin, C. Y., Ouyang, M., & Guo, Y. K. (2013). Biocloud: Cloud computing for biological, genomics, and drug design. *BioMed Research International*, *909470*. doi:10.1155/2013/909470

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Iltifat, H. (2010). Gyromaniac app for iPhone could help gastroenterologists when learning how to perform a colonoscopy. *iMedicalApps.* Retrieved 20 Aug., 2017, from https://www.imedicalapps.com/2010/06/gastroenterologists-iphone-4-medical-spatial-orientation/#

Itani, W., Kayssi, A., & Chehab, A. (2016). Wireless Body Sensor Networks: Security, Privacy, and Energy Efficiency in the Era of Cloud Computing. *International Journal of Reliable and Quality E-Healthcare*, *5*(2), 1–30. doi:10.4018/IJRQEH.2016040101

Janevski, T. (2009, January 10-13). 5G Mobile Phone Concept. *Paper presented at the 2009 6th IEEE Consumer Communications and Networking Conference*.

Johnson, P. T., Zimmerman, S. L., Heath, D., Eng, J., Horton, K. M., Scott, W. W., & Fishman, E. K. (2012). The iPad as a mobile device for CT display and interpretation: Diagnostic accuracy for identification of pulmonary embolism. *Emergency Radiology*, *19*(4), 323–327. doi:10.100710140-012-1037-0 PMID:22450843

Kastania, A. N., & Moumtzoglou, A. (2012). Quality Implications of the Medical Applications for 4G Mobile Phones. *International Journal of Reliable and Quality E-Healthcare*, *1*(1), 58–67. doi:10.4018/ijrqeh.2012010106

Kindig, D. A. (2007). Understanding population health terminology. *The Milbank Quarterly*, *85*(1), 139–161. doi:10.1111/j.1468-0009.2007.00479.x PMID:17319809

Kiselev, A. R., Gridnev, V. I., Shvartz, V. A., Posnenkova, O. M., & Dovgalevsky, P. Y. (2012). Active ambulatory care management supported by short message services and mobile phone technology in patients with arterial hypertension. *Journal of the American Society of Hypertension*, *6*(5), 346–355. doi:10.1016/j.jash.2012.08.001 PMID:22995803

Kitanov, S., Monteiro, E., & Janevski, T. (2016, April 18-20). 5G and the Fog — Survey of related technologies and research directions. *Paper presented at the 2016 18th Mediterranean Electrotechnical Conference (MELECON)*.

Koumpouros, Y. (2015). Big Data in Healthcare. In A. Information Resources Management (Ed.), Healthcare Administration: Concepts, Methodologies, Tools, and Applications (pp. 23-46). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-6339-8.ch002

Kuo, A. M. (2011). Opportunities and challenges of cloud computing to improve health care services. *Journal of Medical Internet Research*, *13*(3), e67. doi:10.2196/jmir.1867 PMID:21937354

Lupse, O. S., Vida, M., & Stoicu-Tivadar, L. (2012). Cloud computing technology applied in healthcare for developing large scale flexible solutions. *Studies in Health Technology and Informatics*, *174*, 94–99. PMID:22491119

Market Research Future. (2017). Global Big Data in Healthcare Market Promises a Striking Revenue of USD 35 Billion By 2021. Retrieved 27 Aug., 2017, from http://www.prnewswire.com/news-releases/global-big-data-in-healthcare-market-promises-a-striking-revenue-of-usd-35-billion-by-2021-300472980.html

Markets and Markets. (2017). Healthcare Analytics Market worth 24.55 Billion USD by 2021. Retrieved 25 Aug., 2017, from http://www.marketsandmarkets.com/PressReleases/healthcare-data-analytics.asp

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Martin, T. (2012). Assessing mHealth: Opportunities and barriers to patient engagement. *Journal of Health Care for the Poor and Underserved*, *23*(3), 935–941. doi:10.1353/hpu.2012.0087 PMID:24212144

Masalski, M., Kipiński, L., Grysiński, T., & Kręcicki, T. (2016). Hearing Tests on Mobile Devices: Evaluation of the Reference Sound Level by Means of Biological Calibration. *Journal of Medical Internet Research*, *18*(5), e130. doi:10.2196/jmir.4987 PMID:27241793

Medical design. (2013). Platform makes possible remote fetal monitoring and cardiac care. Retrieved 10 Aug., 2017, from http://medicaldesign.com/business/platform-makes-possible-remote-fetal-monitoring-and-cardiac-care

Metelmann, C., & Metelmann, B. (2017). Can a Mobile Video Communication Improve Patient Safety in Prehospital Emergency Medicine? *International Journal of Reliable and Quality E-Healthcare*, *6*(4), 27–36. doi:10.4018/IJRQEH.2017100104

Mirza, H., & El-Masri, S. (2013). National Electronic Medical Records integration on Cloud Computing System. *Studies in Health Technology and Informatics*, *192*, 1219. PMID:23920993

Mobisante Inc. (2017). MobiSante Inc. web site. Retrieved 19 Aug., 2017, from http://www.mobisante.com

Moumtzoglou, A. (2014). Demystifying Quality of Healthcare in the Cloud. In M. Anastasius & N. K. Anastasia (Eds.), *Cloud Computing Applications for Quality Health Care Delivery* (pp. 284–301). Hershey, PA, USA: IGI Global. doi:10.4018/978-1-4666-6118-9.ch015

Moumtzoglou, A., & Kastania, A. (Eds.). (2014). *Cloud computing applications for quality health care delivery*. Hershey, PA: IGI Global. doi:10.4018/978-1-4666-6118-9

Moumtzoglou, A., & Pouliakis, A. (2015). Mapping Population Health Management Roadmap into Cervical Cancer Screening Programs. *International Journal of Reliable and Quality E-Healthcare*, *4*(3), 1–18. doi:10.4018/IJRQEH.2015070101

O'Connor, Y., Heavin, C., & O'Donoghue, J. (2015). Review of Key Stakeholders for an mHealth Pilot Study in Malawi Motivations and Expectations. *International Journal of Reliable and Quality E-Healthcare*, *4*(2), 51–66. doi:10.4018/IJRQEH.2015040105

Pandher, P. S., & Bhullar, K. K. (2014). Smartphone applications for seizure management. *Health Informatics Journal*. doi:10.1177/1460458214540906 PMID:25038202

Patel, R. P. (2012). Cloud computing and virtualization technology in radiology. *Clinical Radiology*, *67*(11), 1095–1100. doi:10.1016/j.crad.2012.03.010 PMID:22607956

Patrick, K., Raab, F., Adams, M. A., Dillon, L., Zabinski, M., Rock, C. L., ... Norman, G. J. (2009). A text message-based intervention for weight loss: Randomized controlled trial. *Journal of Medical Internet Research*, *11*(1), e1. doi:10.2196/jmir.1100 PMID:19141433

Pavlopoulos, S., Kyriacou, E., Berler, A., Dembeyiotis, S., & Koutsouris, D. (1998). A novel emergency telemedicine system based on wireless communication technology--AMBULANCE. *IEEE Transactions on Information Technology in Biomedicine*, *2*(4), 261–267. doi:10.1109/4233.737581 PMID:10719536

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Pouliakis, A., Archondakis, S., Karakitsou, E., & Karakitsos, P. (2014). Cloud Computing for Cytopathologists. In A. Moumtzoglou & A. Kastania (Eds.), *Cloud computing applications for quality health care delivery* (pp. 250–271). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-6118-9.ch013

Pouliakis, A., Archondakis, S., Margari, N., & Karakitsos, P. (2016). From telecytology to mobile cytopathology: Past, present and future. In A. Moumtzoglou (Ed.), *M-Health Innovations for Patient-Centered Care*. Hershey, PA: IGI Global. doi:10.4018/978-1-4666-9861-1.ch012

Pouliakis, A., Spathis, A., Kottaridi, C., Mourtzikou, A., Stamouli, M., Archondakis, S., ... Karakitsos, P. (2014). Cloud Computing for BioLabs. In A. Moumtzoglou & A. Kastania (Eds.), *Cloud computing applications for quality health care delivery* (pp. 228–249). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-6118-9.ch012

Quinn, C. C., Shardell, M. D., Terrin, M. L., Barr, E. A., Ballew, S. H., & Gruber-Baldini, A. L. (2011). Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care*, *34*(9), 1934–1942. doi:10.2337/dc11-0366 PMID:21788632

Rana, A. (2017). The Immense Potential of M-Care in India: Catering Better to Patients Needs in the Context of a Fragmented Healthcare System. *International Journal of Reliable and Quality E-Healthcare*, *6*(4), 1–3. doi:10.4018/IJRQEH.2017100101

Ribu, L., Holmen, H., Torbjornsen, A., Wahl, A. K., Grottland, A., Smastuen, M. C., ... Arsand, E. (2013). Low-intensity self-management intervention for persons with type 2 diabetes using a mobile phone-based diabetes diary, with and without health counseling and motivational interviewing: Protocol for a randomized controlled trial. *JMIR Research Protocols*, *2*(2), e34. doi:10.2196/resprot.2768 PMID:23978690

Rosales Saurer, B., Mueller-Gorchs, M., & Kunze, C. (2009). Scenario-based design of an ICT platform for mobile information services in ambulatory care nursing. *Studies in Health Technology and Informatics*, *146*, 64–68. PMID:19592810

Rosenthal, A., Mork, P., Li, M. H., Stanford, J., Koester, D., & Reynolds, P. (2010). Cloud computing: A new business paradigm for biomedical information sharing. *Journal of Biomedical Informatics*, *43*(2), 342–353. doi:10.1016/j.jbi.2009.08.014 PMID:19715773

Roussos, G., Marsh, A. J., & Maglavera, S. (2005). Enabling pervasive computing with smart phones. *IEEE Pervasive Computing*, *4*(2), 20–27. doi:10.1109/MPRV.2005.30

Sama, P. R., Eapen, Z. J., Weinfurt, K. P., Shah, B. R., & Schulman, K. A. (2014). An evaluation of mobile health application tools. *JMIR mHealth and uHealth*, *2*(2), e19. doi:10.2196/mhealth.3088 PMID:25099179

Skrovseth, S. O., Arsand, E., Godtliebsen, F., & Hartvigsen, G. (2012). Mobile phone-based pattern recognition and data analysis for patients with type 1 diabetes. *Diabetes Technology & Therapeutics*, *14*(12), 1098–1104. doi:10.1089/dia.2012.0160 PMID:23035775

Sniderman, A. D., LaChapelle, K. J., Rachon, N. A., & Furberg, C. D. (2013). The necessity for clinical reasoning in the era of evidence-based medicine. *Mayo Clinic Proceedings*, *88*(10), 1108–1114. doi:10.1016/j.mayocp.2013.07.012 PMID:24079680

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Software, M. I. M. (2017). MIM Software web site. Retrieved August 20, 2017, from https://www.mimsoftware.com

Soman, A. K. (2011). Cloud-based solutions for healthcare IT. Enfield, NH: Science Publishers; Distributed by CRC Press. doi:10.1201/b10737

Spat, S., Holl, B., Petritsch, G., Schaupp, L., Beck, P., & Pieber, T. R. (2013). Automatic system testing of a decision support system for insulin dosing using Google Android. *Studies in Health Technology and Informatics*, *186*, 187–191. PMID:23542995

Spyra, G., Buchanan, W. J., Cruickshank, P., & Ekonomou, E. (2014). Cloud-Based Identity and Identity Meta-Data: Secure and Control of Data in Globalization Era. [IJRQEH]. *International Journal of Reliable and Quality E-Healthcare*, *3*(1), 49–66. doi:10.4018/ijrqeh.2014010105

Srirama, S. (2006). Publishing and discovery of mobile web services in peer to peer networks. *Paper presented at the First International Workshop on Mobile Services and Personalized Environments*.

Srirama, S., Jarke, M., & Prinz, W. (2007). Mobile web service discovery in peer to peer networks. *Paper presented at the 5th International Workshop on Ubiquitous Mobile Information and Collaboration Systems*.

Srirama, S., Jarke, M., Zhu, H., & Prinz, W. (2008, June 8-13). Scalable Mobile Web Service Discovery in Peer to Peer Networks. *Paper presented at the Third International Conference on Internet and Web Applications and Services*. 10.1109/ICIW.2008.52

Su, C. J., & Chu, T. W. (2014). A mobile multi-agent information system for ubiquitous fetal monitoring. *International Journal of Environmental Research and Public Health*, *11*(1), 600–625. doi:10.3390/ijerph110100600 PMID:24452256

Subversus Interactive. (2011). Gyromaniac. Retrieved 16 Aug., 2017, from http://subvers.us

Toomey, R. J., Ryan, J. T., McEntee, M. F., Evanoff, M. G., Chakraborty, D. P., McNulty, J. P., ... Brennan, P. C. (2010). Diagnostic efficacy of handheld devices for emergency radiologic consultation. *AJR. American Journal of Roentgenology*, *194*(2), 469–474. doi:10.2214/AJR.09.3418 PMID:20093611

Topol, E. J. (2012). *The creative destruction of medicine: how the digital revolution will create better health care*. New York: Basic Books.

Toye, E., Sharp, R., Madhavapeddy, A., Scott, D., Upton, E., & Blackwell, A. (2007). Interacting with mobile services: An evaluation of camera-phones and visual tags. *Personal and Ubiquitous Computing*, *11*(2), 97–106. doi:10.100700779-006-0064-9

University of Illinois College of Engineering. (2017). New handheld spectral analyzer uses power of smartphone to detect disease. Retrieved 10 Aug., 2017, from https://www.eurekalert.org/pub_releases/2017-08/uoic-nhs081017.php

Vaisanen, O., Makijarvi, M., & Silfvast, T. (2003). Prehospital ECG transmission: Comparison of advanced mobile phone and facsimile devices in an urban Emergency Medical Service System. *Resuscitation*, *57*(2), 179–185. doi:10.1016/S0300-9572(03)00028-5 PMID:12745186

Wavelet Health. (2017). Wavelet Health web site. Retrieved 14 Aug., 2017, from https://wavelethealth.com

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Waxer, N., Ninan, D., Ma, A., & Dominguez, N. (2013). How cloud computing and social media are changing the face of health care. *The Physician Executive Journal, 39*(2), 58-60, 62.

Webb, G. (2012). Making the cloud work for healthcare: Cloud computing offers incredible opportunities to improve healthcare, reduce costs and accelerate ability to adopt new IT services. *Health Management Technology*, *33*(2), 8–9. PMID:22397112

Wilson, J. (2006). 3G to Web 2.0? Can Mobile Telephony Become an Architecture of Participation? *Convergence*, *12*(2), 229–242. doi:10.1177/1354856506066122

Witzke, K., & Specht, O. (2017). M-Health Telemedicine and Telepresence in Oral and Maxillofacial Surgery: An Innovative Prehospital Healthcare Concept in Structurally Weak Areas. *International Journal of Reliable and Quality E-Healthcare*, *6*(4), 37–48. doi:10.4018/IJRQEH.2017100105

Ybarra, M. L., Holtrop, J. S., Prescott, T. L., & Strong, D. (2014). Process evaluation of a mHealth program: Lessons learned from Stop My Smoking USA, a text messaging-based smoking cessation program for young adults. *Patient Education and Counseling*, *97*(2), 239–243. doi:10.1016/j.pec.2014.07.009 PMID:25103183

Zerth, J., Besser, J., & Reichert, A. (2012). Effectiveness and efficiency for ambulatory care assisted by mobile technological devices. *Biomedizinische Technik. Biomedical Engineering*, *57*(Suppl. 1.). doi:10.1515/bmt-2012-4062 PMID:23096280

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 7(2); edited by Anastasius Moumtzoglou; pages 1-26; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 2

Development and Design Methodologies

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 5
A Yes/No Answer Generator Based on Sentiment-Word Scores in Biomedical Question Answering

Mourad Sarrouti
Sidi Mohammed Ben Abdellah University, Morocco

Said Ouatik El Alaoui
Sidi Mohammed Ben Abdellah University, Morocco

ABSTRACT

Background and Objective: Yes/no question answering (QA) in open-domain is a longstanding challenge widely studied over the last decades. However, it still requires further efforts in the biomedical domain. Yes/no QA aims at answering yes/no questions, which are seeking for a clear "yes" or "no" answer. In this paper, we present a novel yes/no answer generator based on sentiment-word scores in biomedical QA. Methods: In the proposed method, we first use the Stanford CoreNLP for tokenization and part-of-speech tagging all relevant passages to a given yes/no question. We then assign a sentiment score based on SentiWordNet to each word of the passages. Finally, the decision on either the answers "yes" or "no" is based on the obtained sentiment-passages score: "yes" for a positive final sentiment-passages score and "no" for a negative one. Results: Experimental evaluations performed on BioASQ collections show that the proposed method is more effective as compared with the current state-of-the-art method, and significantly outperforms it by an average of 15.68% in terms of accuracy.

INTRODUCTION

The large size of literature in the biomedical domain makes it difficult for information seekers even in their field of interest to find the information they need (McDermid, Kristjanson, & Spry, 2010). The most used for accessing to biomedical information are information retrieval (IR) systems, such as

DOI: 10.4018/978-1-7998-1204-3.ch005

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

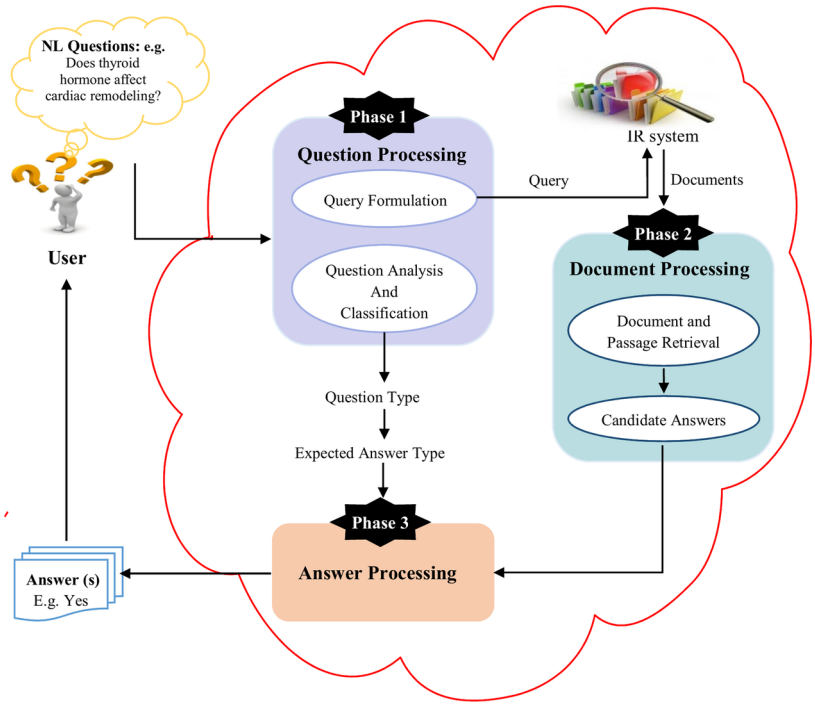
PubMed[1] which gives access to the MEDLINE[2] biomedical bibliographic database (Hristovski, Dinevski, Kastrin, & Rindflesch, 2015). Indeed, finding sufficient and short precise answers is a challenging task for classical IR systems (Wren, 2011). In addition, in classical IR systems, the users have often to deal with the burden of studying and filtering the returned citations of their queries so as to find the precise information they were looking for. Therefore, to minimize searching and browsing time while maximizing the usefulness of that knowledge is a growing interest for biomedical question answering systems (Bauer & Berleant, 2012). Question answering (QA) regards a sophisticated form of IR characterized by information needs that are expressed as natural language statements or questions (Wren, 2011). It aims at providing inquiries with specific pieces of information as an answer, by automatically analyzing thousands of articles, ideally, in less than a few seconds. Typically, an automated QA system consists of three main processing phases, which can be studied and developed independently (Athenikos & Han, 2010; Cao et al., 2010; Neves & Leser, 2015): (1) question processing, (2) document processing, and (3) answer processing. Figure 1 illustrates the generic architecture of a biomedical QA system.

Given an input biomedical question, the question is first handed over to the question processing phase. The latter consists of the following components: (a) question analysis for extracting some useful information such as biomedical entity names, and semantic relationships; (b) question classification for identifying the answer format and the topic (Cao et al., 2010; Patrick & Li, 2012; Roberts et al., 2014; Lopes et al., 2014; Sarrouti et al., 2015); (c) query formulation for constructing IR-style query by transforming the question into a canonical form. The output of this phase is an appropriate query which is used as input to document processing, the second phase. An IR system is normally used to retrieve the relevant documents (Sarrouti & Alaoui, 2016). Then, passages are extracted which serve as answer candidates as well as an input to the last phase, answer processing; in this phase, the system uses an appropriate answer extraction algorithm to estimate the qualities of the candidate answers. Finally, the top-ranked candidate answers and the raw texts from which the answers were extracted are shown to the user (Sarrouti & Ouatik, 2017).

In spite of the importance of the process, yes/no-style QA has not been intensively studied in the biomedical domain, compared to other types of QA, such as the factoid-style QA (Weissenborn et al., 2013; Yang et al., 2015; Hristovski et al., 2015) and the definition-style QA (Lee et al., 2006; Yu et al., 2007; Cao et al., 2011). Yes/no QA aims at answering yes/no questions, which are seeking for a clear "yes" or "no" answer (Green & Carberry, 1994). For instance, a biomedical yes/no QA system would provide "yes" as an answer for the yes/no question "Does the histidine-rich Ca-binding protein (HRC) interact with triadin?". Even though there are only two possible answers, "yes" or "no," such questions can be quite hard to answer due to the complicated sentiment analysis process of the candidate answers. This challenging task has motivated the recent launch of the BioASQ[3] challenge: an EU-funded support action to set up challenges on biomedical semantic indexing and QA (Tsatsaronis et al., 2015). The BioASQ challenge consists of two main tasks, i.e., task a: large-scale online biomedical semantic indexing and task b: biomedical semantic QA. The major goal of task b is to assess the performance of participating systems in different stages of the QA process. Moreover, task b is sub-divided into two phases: phase A and phase B. Indeed, in phase B, BioASQ provided researchers with biomedical questions and their respective question types (i.e., yes/no, factoid, list, summary) as well as gold-standard information for phase A, such as manually curated relevant concepts, documents, passages and RDF triples (Balikas et al., 2013). The participants were requested to provide an exact answer: "yes" or "no" for yes/no questions, a short paragraph for summary questions, and a single or a list of short answers for factoid and list questions, respectively.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A typical biomedical question answering system architecture



In spite of the launch of the biomedical QA track at the BioASQ challenge and the great effort of biomedical experts to create benchmark datasets, only few answer extraction methods for yes/no questions have been presented in the biomedical domain, compared to other question types such as factoid, list, and summary. The most recent yes/no biomedical answer extraction method is the one which was presented by Neves (2015). The author has made a decision on either the answer is "yes" or "no," based on the sentiment analysis predictions provided by the in-memory based database (IMDB) technology (Plattner, 2013). The IMDB returns 10 types of sentiments, namely "strong positive sentiment," "strong positive emoticon," "weak positive sentiment," "weak positive emoticon," "strong negative sentiment," "strong negative emoticon," "major problem," "weak negative sentiment," "weak negative emoticon" and "minor problem". Indeed, the author has grouped these sentiments into four coarse categories, i.e., "positive strong," "positive weak," "negative strong," "negative weak," and then into the two main sentiments "positive" or "negative" and decision for the answers "yes" or "no" was based on these two coarse sentiments.

Although the previous method has proven to be quite successful at answering yes/no questions in biomedical QA, yes/no answer extraction system still requires further efforts in order to improve its performance. For this purpose, we propose in this paper a yes/no answer extraction method based on sentiment-word scores. We first use the Stanford CoreNLP (Manning, Surdeanu, Bauer, Finkel, Bethard, & Closky, 2014) tool for tokenization and part-of-speech tagging all passages for a given yes/no question. We then assign to each word in passages a sentiment score based on SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010): a publicly available lexical resource for sentiment analysis and opinion mining. Finally, the decision for the answers "yes" or "no" is based on the final sentiment score which is the sum of all word scores: "yes" for a positive final sentiment passages score and "no" for a negative

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

one. Extensive experiments on BioASQ collections demonstrate the effectiveness of our proposed yes/no answer extraction method. The results show that our method is more efficient as compared with the current state-of-the-art method presented in (Neves, 2015) and outperforms it by an average of 15.68% in terms of accuracy.

The remainder of the paper is organized as follows. In Section 2, we introduce related work and discussion about the main yes/no answer extraction approaches. Section 3 describes the overall architecture of the proposed yes/no answer extraction method. Several comprehensive experiments are performed to evaluate the effectiveness of our approach in Section 4 where experimental settings, evaluation metrics, benchmark datasets, and results are presented. Finally, Section 5 concludes this paper and presents directions for future research.

RELATED WORK

The task of automatically answering natural language questions has been intensively studied, especially since the introduction of QA Track in the Text REtrieval Conference (TREC[4]) in 1999, and the presentation of biomedical QA in the BioASQ challenge (Tsatsaronis et al., 2015). The TREC QA challenge deals only with factoid, list and definitional questions, which were not restricted to any domain. The BioASQ challenge differs from the aforementioned challenge in two main points. First, the questions of the challenge are restricted to the biomedical domain. Second, it includes two new types of questions, namely yes/no questions and summary questions.

In spite of the importance of yes/no biomedical QA system, only few studies have been presented compared to other types of QA, such as the factoid-style QA (Weissenborn et al., 2013; Yang et al., 2015; Hristovski et al., 2015) and the definition-style QA (Lee et al., 2006; Yu et al., 2007; Cao et al., 2011). Moreover, the majority of the proposed yes/no QA systems deal with open-domain yes/no questions, which are not restricted to any domain. Yes/no QA (Green & Carberry, 1994) aims at answering yes/no questions, which are seeking for a clear "yes" or "no" answer. In this direction, Kanayama et al. (2012) have investigated a solution to yes/no QA in the open domain by leveraging the power of information sources accessed through a factoid-style QA system. By aggregating the answers and confidence values from a factoid-style QA system, the authors have determined the correctness of the entire proposition or the substitutions that make the proposition false. He and Dai (2011) have designed a model to score the relevance of the answers and the questions, and a feature function combining the relevance and opinion scores to classify each answer to be "yes" or "no". The authors first have proposed a model to compute the relevance of asymmetric documents and used it as one of the features to give a weight for each sentence in the answers. Then, they have introduced three features, including the relevance feature, the sentiment-word score feature and the position weight feature for classifying the answers to general questions. Finally, they have integrated these three features to find a new feature function that classifies the answers into three types: "yes," "no," and "neutral". Bdour and Gharaibeh (2013) have proposed a formal model for a lightweight semantic-based open-domain yes/no Arabic QA system based on passage retrieval. The authors have shown that using a semantic logical representation based on query expansion techniques frequently improve the precision of the system. Their experimental results demonstrated that employing passage retrieval achieves a better precision by retrieving more passages that contain relevant answers to the questions. Kim et al. (2014) have developed a QA approach to answering yes/no questions relevant to civil laws in legal bar exams. In their proposed system, the authors first have identified legal

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

documents relevant to the exam questions, and then extracted the final answers of the questions by analyzing the relevant documents. They have constructed a knowledge base by analyzing negation patterns and antonyms in the civil law articles. In addition, the authors have extracted deep linguistic features with lexical, syntactic information based on morphological analysis and dependency trees, and lexical semantic information using the Kadokawa thesaurus. Furthermore, they have classified the questions into easy and difficult ones and proposed the two-phase method for answering yes/no questions. They have answered easy questions by negation/antonym detection. For more difficult questions, they have adapted a machine-learning method based on morphological, syntactic, and lexical semantic analysis as features. Abacha and Zweigenbaum (2015) have developed a medical QA system, namely MEANS, based on natural language processing techniques and semantic Web technologies. MEANS consists of three main phases: (1) corpora annotation, (2) question analysis, and (3) answer search. They have proposed a method for answer search based on semantic search and query relaxation. Neves (2015) has proposed a biomedical yes/no answer extraction method based on the sentiment analysis predictions provided by the IMDB technology (Plattner, 2013). As it has already been stated before, the author has grouped the 10 types of sentiments returned by the IMBD into the two main sentiments, "positive" or "negative," in order to decide whether the answer should be "yes" or "no."

The previous biomedical yes/no answer extraction method presented by Neves (2015), which was evaluated on benchmark datasets provided by the BioASQ challenge, has proven to be quite successful at answering yes/no questions in biomedical QA. However, yes/no answer extraction component in biomedical QA still requires further efforts in order to improve its performance. Hence, this paper tackles yes/no QA in the biomedical domain by exploiting sentiment-word scores provided by the SentiWordNet lexical resource. The latter, which is based on lexical approaches, has been widely used in sentiment analysis and opinion mining methods presented in social networks' text mining research area (Abbasi et al., 2011; Paltoglou & Thelwall, 2012; Mostafa, 2013; Thelwall & Buckley, 2013). The key idea behind using the SentiWordNet word score in the proposed yes/no QA is that all passages are somehow related to the yes/no question and that computing sentiment scores of these passages could be used to distinguish between the two possible answers, i.e., "yes" and "no". The contributions that we make in this paper can be summarized in the following points:

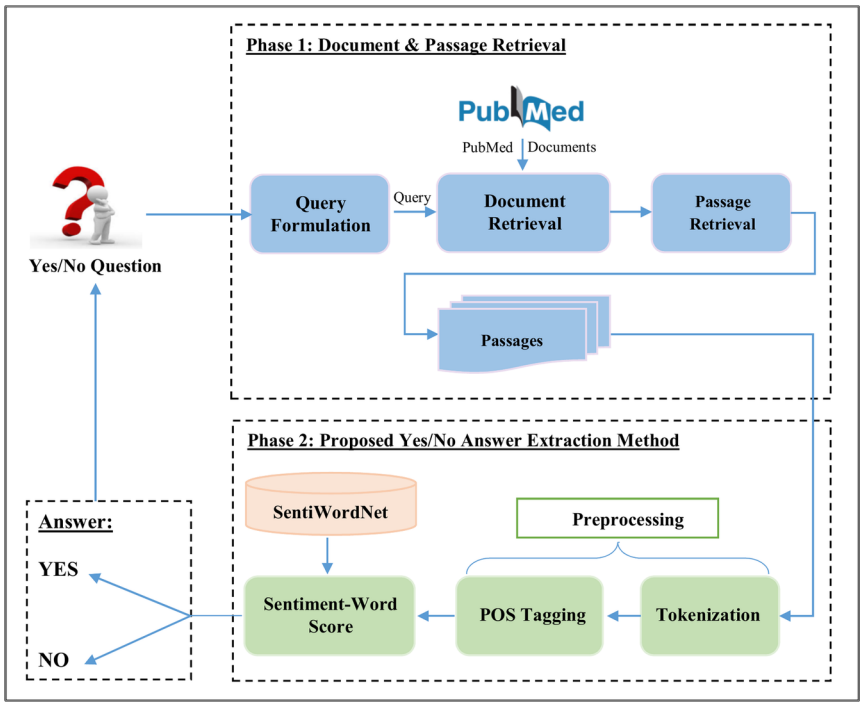
1. As far as we know, there are no studies in the literature that use the SentiWordNet lexical resource for answering yes/no questions;
2. We propose a new yes/no answer extraction method in biomedical QA based on the sentiment word scores provided by the SentiWordNet lexical resource;
3. We conduct experiments on benchmark datasets, demonstrating that the proposed method significantly outperforms the current state-of-the-art method presented in (Neves, 2015).

METHODS

In this section, we present the proposed yes/no answer extraction method, which is comprised of two stages: document & passage retrieval and yes/no answer generator. The flowchart of the proposed method is shown in Figure 2. The two phases are described sequentially in the following subsections.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. System overview of the proposed yes/no answer extraction method



Document and Passage Retrieval

In general, an IR (Sarrouti & Alaoui, 2016) system is used for document and passage retrieval to a given question in QA. However, in this work, we used the golden (correct) documents and passages from benchmark datasets provided by the BioASQ challenge (Tsatsaronis et al., 2015). BioASQ realized a set of biomedical questions with their golden documents and passages in order to evaluate the answer extraction methods in their best way.

Proposed Yes/No Answer Extraction Method

After retrieving the n golden passages $\{p_1, p_2, p_3, ..., p_n\}$ to a given yes/no question from benchmark datasets, we forward them one by one to the Stanford CoreNLP (Manning et al., 2014) for tokenization and part-of-speech tagging in order to extract a set of words accompanied by their POS tags. The Stanford CoreNLP was used since recent research shows a good performance of the latter in the biomedical domain (Lin & Wilbur, 2007). We then compute the sentiment score for each passage to see whether the answer should be "yes" or "no". Assuming that there are k words (w) in a passage p, the final sentiment score (SC) of the passage is defined by the following equation:

$$SC(p) = \sum_{i=1}^{k} \text{SentiWordNet}\ (\text{w}_\text{i}) \tag{1}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The sentiment score for each word is obtained using SentiWordNet[5] 3.0: a lexical resource for sentiment analysis and opinion mining (Baccianella et al., 2010). The latter assigns to each synset of WordNet three sentiment scores: "positivity," "negativity," and "objectivity". Indeed, the key idea behind using the SentiWordNet 3.0 lexical resource is that it is the result of the automatic annotation of all the synsets of WordNet 3.0 according to the notions of "positivity," "negativity," and "neutrality" (Baccianella et al., 2010). The WordNet, which is a large lexical database of English, is comprised of 155,287 words and 117,659 synsets, also called synonyms (George, 1995). On the other hand, March et al. (2013) have shown that SentiWordNet 3.0 outperforms other sentiment lexicons in the determination of the polarity, such as Bing Liu's Opinion Lexicon (Hu & Liu, 2004) and MPQA Subjectivity Lexicon (Wilson, Wiebe, & Hoffmann, 2005).

Algorithm 1. *Yes/No answer generator*

```
Input ← yes/no question Q and set of golden passages P
Output ← answer: "yes" or "no"
postive ← 0 » number of positive passage
negative ← 0 » number of negative passage
i ← 1
function PreProcessing (p)
TOKEN [1...m] ← TokenizationAndPOSTagging (p)
return TOKEN
end function
do
W [1...m] ← PreProcessing (P[i]) » extract a set of words and their POS
tags of a passage p
score ← 0.0
j ← 1
do
score ← score + SentiWordNet (W[j])
j ← j + 1
while j ≤ m » m is the size of a set of words W
if score ≥ 0 then
positive ← positive + 1
else
negative ← negative + 1
end if
i ← i + 1
while i ≤ n » n is the size of a set of passages P
if positive ≥ negative then » the decision on either the answers "yes" or
"no" is based on the number positive and negative passages
output ← "yes"
else
output ← "no"
end if
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In Algorithm 1, we summarize all steps of the proposed yes/no answer extraction method in biomedical QA. As stated earlier, our approach first takes all golden passages for a given yes/no question. After that, terms and their POS tags of passages are extracted in the preprocessing step. Then, it computes the sentiment score for each passage based on the SentiWordNet-word scores. Finally, the decision on either the answers "yes" or "no" is based on the number of positive and negative passages: "yes" for a positive final sentiment-passages score and "no" for a negative one. Figure 3 shows an example of the whole process, i.e., tokenization, part-of-speech tagging, and sentiment score assignment for one passage of the yes/no question "Does the CTCF protein colocalize with cohesin?"

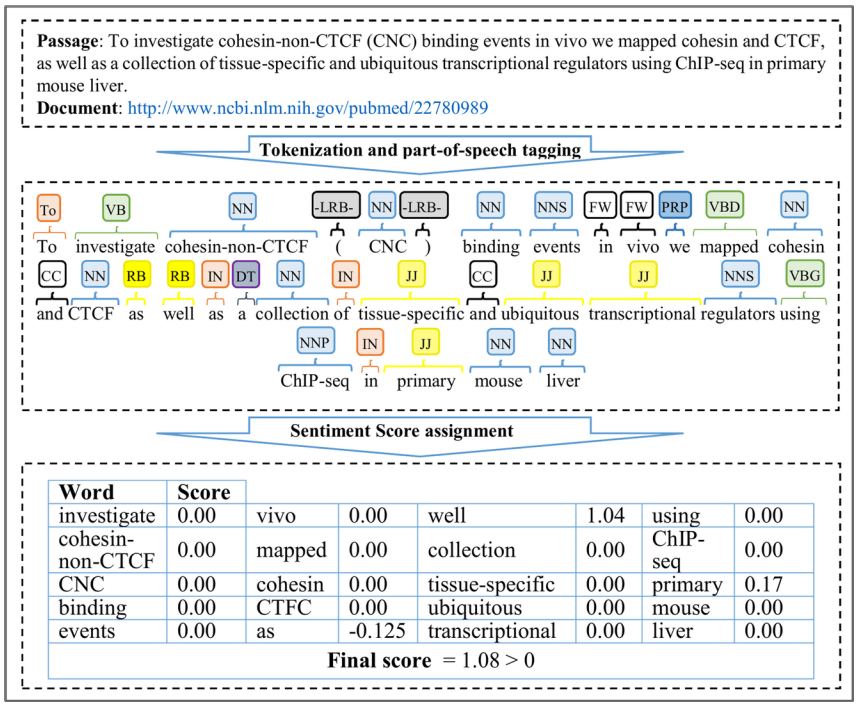
EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we first report our pilot evaluation of the proposed biomedical yes/no answer extraction method on benchmark datasets, provided by the BioASQ challenge. We then compare our method with the current state-of-the-art method. We finally show an error case study to understand the source of errors in yes/no answer extraction.

Datasets

To validate the efficiency of the proposed yes/no answer extraction method in biomedical QA, we used the publicly available benchmark datasets provided by the BioASQ challenge (Tsatsaronis et al., 2015).

Figure 3. An example of tokenization, part-of-speech tagging and sentiment score assignment



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The challenge, within the 2015 edition, realized a training dataset of 810 biomedical questions-answers pairs and five batches of testing data which are used as test sets to evaluate the participating systems in task b. There were four types of questions: "yes/no," "factoid," "list," and "summary" questions (Balikas et al., 2013). Table 1 shows these five batches of testing datasets and the number of questions assigned to each type.

Furthermore, in phase B of task b, the released biomedical questions were accompanied by the golden (correct) articles as well as passages from benchmark datasets (Balikas et al., 2013). The participating systems had to respond with exact answers (e.g., "yes" or "no" in the case of yes/no questions, named entities in the case of factoid questions, etc.) and ideal answers (paragraph-sized summaries). The following are some examples of yes/no biomedical questions from the BioASQ training dataset:

- Does magnesium sulfate improve outcomes of subarachnoid hemorrhage patients?
- Can NXY-059 be used for treatment of acute ischemic stroke patients?
- Is there an increased risk for cancer in Dyskeratosis Congenita?

Evaluation Metrics

Accuracy is the main evaluation metric used to evaluate the performance of yes/no answer extraction methods in the BioASQ challenge since yes/no QA problem is considered as a binary classification problem (Tsatsaronis et al., 2015). For each yes/no question, the exact answer of yes/no QA system had to be either "yes" or "no." Therefore, the response was compared against the golden exact answer (again "yes" or "no") realized by the BioASQ team of biomedical experts (Balikas et al., 2013). Assuming that there are n yes/no biomedical questions, accuracy is defined in the following equation, where k is the number of correctly answered yes/no questions:

$$\text{Accuracy} = \frac{k}{n} \tag{2}$$

Table 1. The five batches of testing datasets of biomedical questions provided by BioASQ 2015 and the number of questions assigned to each type

Testing Datasets	Question Types				#Questions
	Yes/No	Factoid	List	Summary	
Batch 1	33	26	22	19	100
Batch 2	16	32	28	24	100
Batch 3	29	26	17	28	100
Batch 4	25	29	23	20	97
Batch 5	28	22	24	26	100

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

RESULTS AND DISCUSSION

Although yes/no answer generator has been widely investigated in open-domain QA, few approaches in the biomedical domain are currently able to answer yes/no questions. In this study, our goal was to develop an efficient yes/no answer generator in biomedical QA. Indeed, the decision on either the answers "yes" or "no" is based on the final sentiment-passages score provided by the SentiWordNet 3.0 lexical resource: "yes" for a positive final sentiment-passages score and "no" for a negative one. The assumption was that all passages are somehow related to the yes/no question and that computing sentiment scores of these passages could be used to distinguish between the two possible answers. Table 2 presents the experimental results of the proposed yes/no answer extraction method, as well as the comparison in terms of accuracy with the current state-of-the-art method presented in Neves (2015) on five batches of testing datasets, provided by the BioASQ challenge (Tsatsaronis et al., 2015). The proposed method achieves good performances compared with Neves's (2015) method.

Moreover, it can be seen clearly from Table 2 that in all batches of testing datasets, the proposed yes/no answer extraction method achieves performance improvements over the state-of-the-art method presented in (Neves, 2015). Compared with the latter, which employs the sentiment analysis predictions provided by the in-memory based database (IMDB) technology (Plattner, 2013), our method gives better results (an average improvement of 15.68% in terms of accuracy). Moreover, the increase performance is statistically significant (the p-value is < 0.00001, the result is significant at $p < 0.01$).

In summary, from the results and the analysis we can draw a conclusion that our proposed yes/no answer extraction method allows answering yes/no questions in biomedical QA with high accuracy. The high accuracy proves that our method is convincing, and successfully hunts out a large amount of the incorrect answers. As stated in Section 1, biomedical QA systems consist of three main steps, i.e., question analysis, document/passage retrieval and answer extraction. Indeed, each of these three steps has a significant impact on the overall performance of a biomedical QA system. Therefore, in addition to other components, the answer extraction module plays a vital role in QA systems since it is the final step which aims at retrieving the answers from the retrieved passages of a given biomedical question. Consequently, by improving the performance of answer extraction methods, the overall performance of biomedical QA systems will inevitably improve too.

Table 2. The overall results of the proposed yes/no answer extraction method and the comparison in terms of accuracy with the current state-of-the-art method on five batches of testing data provided by the BioASQ challenge

Models	Accuracy				
	Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
Proposed method	**69.70%**	**62.50%**	**86.21%**	**76.00%**	**60.71%**
Neves (2015)	66.67%	56.25%	62.07%	56.00%	35.71%
Increase performance	+3.03%	+6.25%	+24.14%	+20.00	+25.00%

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Error Analysis

We performed error analysis to understand the source of errors in yes/no answer extraction. Indeed, we found that there are still some mistakes that the proposed yes/no answer extraction method cannot fix. An example is that, when the question is "Does HER2 under-expression lead to favorable response to trastuzumab?," a positive answer might be "Trastuzumab is a monoclonal antibody targeted to the Her2 receptor and approved for treatment of Her2 positive breast cancer," the sentiment score considers this answer to be a positive as there are both "positive" and "approved" in the passage. However, this example should be counted as a "no" answer because trastuzumab is effective only in cancers where Her2 is over-expressed. Dealing with such problems requires more complex semantic analysis, and we may need to parse the passage to get a grammar tree. Nevertheless, parsing and semantic analysis will bring in new errors and make this challenge even more complicated.

CONCLUSION

In this paper, we have proposed a novel yes/no answer generator which employs sentiment-word scores provided by SentiWordNet in biomedical QA systems. Our contributions can be summarized as follows: 1) for a given yes/no biomedical question, we have used the Stanford CoreNLP for tokenization and part-of-speech tagging all golden passages so as to extract a set of words and their POS tags; 2) we have assigned to each word a sentiment score based on the SentiWordNet lexical resource; 3) the decision for the answers, i.e., "yes" or "no" was based on the obtained sentiment-passages scores. Experiments conducted on BioASQ 2015 datasets have demonstrated the effectiveness of the proposed method. The evaluation results show that the proposed method can achieve 15.68% improvement over the current state-of-the-art method in terms of accuracy.

In future research, we intend to improve the proposed system in order to make it able to answer other question types, such as factoid, list, and summary questions, by exploring unified medical language system (UMLS) concepts and more sophisticated algorithms.

ACKNOWLEDGMENT

The authors would like to thank the BioASQ challenge (Tsatsaronis et al., 2015) organizers who provided the datasets used in this study for evaluating the proposed yes/no biomedical answer extraction method.

REFERENCES

Abacha, A. B., & Zweigenbaum, P. (2015). MEANS: A medical question-answering system combining NLP techniques and semantic Web technologies. *Information Processing & Management*, *51*(5), 570–594. doi:10.1016/j.ipm.2015.04.006

Abbasi, A., France, S., Zhang, Z., & Chen, H. (2011). Selecting attributes for sentiment classification using feature relation networks. *IEEE Transactions on Knowledge and Data Engineering*, *23*(3), 447–462. doi:10.1109/TKDE.2010.110

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Athenikos, S. J., & Han, H. (2010). Biomedical question answering: A survey. *Computer Methods and Programs in Biomedicine*, *99*(1), 1–24. doi:10.1016/j.cmpb.2009.10.003 PMID:19913938

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. Proceedings of LREC'10 (Vol. 10, pp. 2200–2204).

Balikas, G., Partalas I., Kosmopoulos, A., Petridis, S., Malakasiotis, P., Pavlopoulos, I., Androutsopoulos, I., Baskiotis, N., Gaussier, E., Artieres, T., & Gallinari, P. (2013). Evaluation Framework Specifications (Project deliverable D4.1, 05/2013).

Bauer, M. A., & Berleant, D. (2012). Usability survey of biomedical question answering systems. *Human Genomics*, *6*(1), 17. doi:10.1186/1479-7364-6-17 PMID:23244628

Bdour, W. N., & Gharaibeh, K. N. (2013). Development of yes/no Arabic question answering system. *International Journal of Artificial Intelligence & Applications*, *4*(1), 51–63. doi:10.5121/ijaia.2013.4105

Cao, Y., Cimino, J. J., Ely, J., & Yu, H. (2010). Automatically extracting information needs from complex clinical questions. *Journal of Biomedical Informatics*, *43*(6), 962–971. doi:10.1016/j.jbi.2010.07.007 PMID:20670693

Cao, Y., Liu, F., Simpson, P., Antieau, L., Bennett, A., Cimino, J. J., ... Yu, H. (2011). AskHERMES: An online question answering system for complex clinical questions. *Journal of Biomedical Informatics*, *44*(2), 277–288. doi:10.1016/j.jbi.2011.01.004 PMID:21256977

George, A. M. (1995). WordNet: A Lexical Database for English. *Communications of the ACM*, *38*(11), 39–41. doi:10.1145/219717.219748

Green, N., & Carberry, S. (1994). Generating indirect answers to yes-no questions. *Proceedings of the Seventh International Workshop on Natural Language Generation INLG '94*. 10.3115/1641417.1641439

He, J., & Dai, D. (2011). Summarization of yes/no questions using a feature function model. *ACML*, *20*, 351–366.

Hristovski, D., Dinevski, D., Kastrin, A., & Rindflesch, T. C. (2015). Biomedical question answering using semantic relations. *BMC Bioinformatics*, *16*(1), 6. doi:10.118612859-014-0365-3 PMID:25592675

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp 168–177).

Kanayama, H., Miyao, Y., & Prager, J. (2012). Answering yes/no questions via question inversion. *Proceedings of COLING*, *2012*, 1377–1392.

Kim, M.-Y., Xu, Y., Goebel, R., & Satoh, K. (2014). Answering yes/no questions in legal bar exams. *Lecture Notes in Computer Science*, *8417*, 199–213. doi:10.1007/978-3-319-10061-6_14

Lee, M., Cimino, J., Zhu, H. R., Sable, C., Shanker, V., Ely, J., & Yu, H. (2006). Beyond information retrieval—medical question answering. *AMIA Annual Symposium Proceeding*, *2006*, 469–473. PMID:17238385

Lin, J., & Wilbur, W. J. (2007). Syntactic sentence compression in the biomedical domain: Facilitating access to related articles. *Information Retrieval*, *10*(4-5), 393–414. doi:10.100710791-007-9029-5

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lopes, C. T., & Ribeiro, C. (2014). Identification and classification of health queries: Co-occurrences vs. domain-specific terminologies. *International Journal of Healthcare Information Systems and Informatics*, *9*(3), 55–71. doi:10.4018/ijhisi.2014070104

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 10.3115/v1/P14-5010

March, M., Ginsca, R. L., Besançon, R., & Mesnard, O. (2013). [LVIC-LIMSI]: Using Syntactic Features and Multi-polarity Words for Sentiment Analysis in Twitter. *Proceedings of the Seventh international workshop on semantic evaluation (SemEval 2013),* Atlanta, Georgia (Vol. 2, pp. 418-424).

McDermid, D. C., Kristjanson, L. J., & Spry, N. (2010). Factors Affecting the Sustainability of Computer Information Systems. *International Journal of Healthcare Information Systems and Informatics*, *5*(1), 1–15. doi:10.4018/jhisi.2010110301

Mostafa, M. M. (2013). More than words: Social networks text mining for consumer brand sentiments. *Expert Systems with Applications*, *40*(10), 4241–4251. doi:10.1016/j.eswa.2013.01.019

Neves, M. (2015). *HPI question answering system in the BioASQ 2015 challenge. In Working Notes for the Conference and Labs of the Evaluation Forum*, Toulouse, France. CLEF.

Neves, M., & Leser, U. (2015). Question answering for biology. *Methods (San Diego, Calif.)*, *74*, 36–46. doi:10.1016/j.ymeth.2014.10.023 PMID:25448292

Paltoglou, G., & Thelwall, M. (2012). Twitter, Myspace, Digg: Unsupervised sentiment analysis in social media. *ACM Transactions on Intelligent Systems and Technology*, *3*(4), 1–19. doi:10.1145/2337542.2337551

Patrick, J., & Li, M. (2012). An ontology for clinical questions about the contents of patient notes. *Journal of Biomedical Informatics*, *45*(2), 292–306. doi:10.1016/j.jbi.2011.11.008 PMID:22142949

Plattner, H. (2014). A Course in in-memory data Management. doi:10.1007/978-3-642-55270-0

Roberts, K., Kilicoglu, H., Fiszman, M., & Demner-Fushman, D. (2014). Automatically classifying question types for consumer health questions. *AMIA Annual Symposium Proceeding*, *2014*, 1018–1027. PMID:25954411

Sarrouti, M., & Alaoui, S. O. E. (2016). A generic document retrieval framework based on UMLS similarity for biomedical question answering system. Smart Innovation, Systems and Technologies, *57*, 207–216. doi:10.1007/978-3-319-39627-9_18

Sarrouti, M., Lachkar, A., & Alaoui, S. O. E. (2015). Biomedical question types classification using syntactic and rule based approach. *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* (Vol. 1, pp 265-272). 10.5220/0005598002650272

Sarrouti, M., & Ouatik, S. E. A. (2017). A passage retrieval method based on probabilistic information retrieval model and UMLS concepts in biomedical question answering. *Journal of Biomedical Informatics*, *68*, 96–103. doi:10.1016/j.jbi.2017.03.001 PMID:28286031

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Thelwall, M., & Buckley, K. (2013). Topic-based sentiment analysis for the social web: The role of mood and issue-related words. *Journal of the American Society for Information Science and Technology*, *64*(8), 1608–1617. doi:10.1002/asi.22872

Tsatsaronis G., Balikas, G., Malakasiotis, P., Partalas, I., Zschunke, M., Alvers, M. R., Weissenborn, D., Krithara, A., Petridis, S., Polychronopoulos, D., Almirantis, Y., Pavlopoulos, J., Baskiotis, N., Gallinari, P., Artieres, T., Ngomo, A. C. N., Heino, N., Gaussier, E., Barrio-Alvers, L., Schroeder, M., Androutsopoulos, I., & Paliouras, G. (2015). An overview of the BIOASQ large-scale biomedical semantic indexing and question answering competition. *BMC Bioinformatics*, *16*(1), 138. doi:10.118612859-015-0564-6 PMID:25925131

Weissenborn, D., Tsatsaronis, G., & Schroeder, M. (2013). Answering factoid questions in the biomedical domain. *BioASQ@CLEF*.

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing* (pp. 347–354). Association for Computational Linguistics. 10.3115/1220575.1220619

Wren, J. D. (2011). Question answering systems in biology and medicine-the time is now. *Bioinformatics (Oxford, England)*, *27*(14), 2025–2026. doi:10.1093/bioinformatics/btr327 PMID:21672971

Yang, Z., Gupta, N., Sun, X., Xu, D., Zhang, C., & Nyberg, E. (2015). *Learning to answer biomedical factoid and list questions: OAQA at BioASQ 3B. Working Notes for the Conference and Labs of the Evaluation Forum*. Toulouse, France: CLEF.

Yu, H., Lee, M., Kaufman, D., Ely, J., Osheroff, J. A., Hripcsak, G., & Cimino, J. J. (2007). Development, implementation, and a cognitive evaluation of a definitional question answering system for physicians. *Journal of Biomedical Informatics*, *40*(3), 236–251. doi:10.1016/j.jbi.2007.03.002 PMID:17462961

ENDNOTES

1 http://www.ncbi.nlm.nih.gov/pubmed/
2 https://www.nlm.nih.gov/pubs/factsheets/medline.html
3 http://bioasq.org/
4 http://trec.nist.gov/
5 http://sentiwordnet.isti.cnr.it/

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 12(3); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 62-74; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 6
Towards an Integrated Electronic Medical Records System for Quality Healthcare in Ghana:
An Exploratory Factor Analysis

Patrick Ohemeng Gyaase
Catholic University College of Ghana, Ghana

Richard Darko-Lartey
Catholic University College of Ghana, Ghana

Harrison William
Catholic University College of Ghana, Ghana

Foster Borkloe
Catholic University College of Ghana, Ghana

ABSTRACT

This article explores the factors militating against the diffusion integration electronic medical records in Ghana. Structured questionnaires were used to collect data from respondents in the healthcare ecosystem and analyzed quantitatively using descriptive and inferential statistics. The findings revealed non-integration of electronic medical records systems among the stakeholders of the healthcare ecosystem mostly due to and interoperability issues. There is also substantial usage of technology in capturing and storing medical records. The key factors militating against the integration of electronic medical records systems are inadequate funds to purchase database management systems, acquisition and maintenance costs of electronic medical records systems, and an unclear return on investment on the integrated electronic medical records systems. The article recommends sustained government financial support, standardization policy and implementation guidelines to facilitate the integration of electronic medical records systems in the various health institutions.

DOI: 10.4018/978-1-7998-1204-3.ch006

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Electronic medical record (EMR) systems is a collection of tools, technologies and process used to create, gather and manage health-related information on an individual that can be consulted by authorized clinicians and staff within the healthcare organization (Tang, Ash, Bates, Overhage, & Sands, 2006). EMR systems are credited with substantial benefits to physicians, clinic practices, and healthcare organizations by facilitating workflow and improving the quality of patients' care and safety (healthit.ahrq.gov, 2015). When integrated among healthcare providers, EMR systems store patient clinical information electronically, enabling instant access to by all providers in the healthcare chain, thereby facilitating provision of coherent and consistent care (Albert & Manda, 2010).

An integrated electronic medical record (IEMR) is becoming the core of computerized healthcare systems. The electronic storage of clinical information creates the potential for computer-based tools to help clinicians significantly enhance the quality of medical care and increase the efficiency of medical practice (Acquah-Swanzy, 2015). These tools may include reminder systems that identify patients who are due for preventative care interventions, alerting systems that detect contraindications among prescribed medications and coding systems that facilitate the selection of correct billing codes for patient encounters (Kurtin, Harrison, Iraca, & Hassan, 2017).

Integrated electronic medical records (IEMR) systems has transformed healthcare worldwide. Some developed countries have built effective IEMR systems throughout the healthcare value chain based on established national standards (Hochheise & Shneiderman, 2011). However, this is not the case in Ghana where individual health institutions have adopted their own EMR systems resulting in interoperability problems and hindering information sharing among healthcare facilities (Acquah-Swanzy, 2015). The improvements in healthcare facilitated by integrated EMR systems has compelled many developing countries to explore its diffusion into their health care delivery systems. Ghana over the years has shown some level of commitments to implement EMR systems in the health care system as evidenced in the various policies and initiatives (MOH, 2005), (Boateng, 2009).

The Ghana Health Service is making effort to implement integrated EMR system through various health sector policies (Adjorlolo & Ellingsen, 2013). In spite of these policies, the extent of EMRS integration in the healthcare delivery system in Ghana leaves much to be desired. Various hospitals continue to use their own EMR systems run on different operating systems with different data and database management systems (Acquah-Swanzy, 2015). This study therefore aims at ascertaining the extent to which hospitals employ electronic medical records systems, examine the extent of EMRs systems integration and key factors influencing the effective deployment and integration of electronic medical records to improve healthcare delivery in Ghana.

INTEGRATED ELECTRONIC MEDICAL RECORDS IN HEALTHCARE

Integrated Electronic Medical Record (EMR) systems are becoming an essential fabric of modern healthcare. EMR systems have evolved from stand-alone single computer record-keeping to an integrated, enterprise-wide system that ensures real time access to accurate patient healthcare data (Wiliam & Michael, 2009). The ease of data retrieval with a click of mouse from integrated EMR systems and real-time access to patients' information by clinicians and other stakeholders result in efficiency gains and quality care delivery.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Integrated EMR systems also encourage collaboration within the healthcare system through dissemination of health information and sharing data in real-time among stakeholders in the healthcare delivery system. Integrated EMR system has the potential to reduce cost of healthcare delivery in terms of manpower required and savings from stationery thus improving efficiency (Hoque, Rakibul, & Golam, 2016).

Implementation of integrated electronic medical records systems has the potential to improve overall quality of care from both clinical, organizational perspectives. From clinical perspective, improvements in the quality of care, reduction in medical errors, and other improvements in patient-level measures have been evidenced in the deployment of integrated EMR systems. From organizational perspectives, improvements in financial and operational performance, satisfaction among patients and improved performance of clinicians are also well documented (Ashish, et al., 2009).

Enabling Environment for Electronic Medical Records Integration

A successful implementation of integrated electronic medical records systems requires certain enablers if the attributable benefits could be realized.

Availability of effective and efficient network connectivity is a critical enabler for successful implementation of integrated EMR systems. This would ensure that EMR system can link all stakeholders in the health care chain such as laboratories, imaging systems, pharmacies, hospitals and clinics. This would facilitate seamless exchange of patients' laboratory results, prescription and other clinical and personal information in an integrated healthcare environment (HealthIT.gov, 2014).

Appreciable levels of ICT competencies among personnel is also essential for the smooth implementation of integrated EMR system in a healthcare system (Khalifa, 2013). Healthcare personnel with exposure to ICT tools in their working environment are likely to be receptive and could easily adapt to an integrated EMR system environment. User resistance due to lack of ICT skills is known to limit the adoption of ICT innovations in organization (Price, 2013). The absence of existing competent users of computers would require further investments in ICT training to ensure successful implementation.

Positive attitude of staff toward the usage of integrated EMR systems is another important factor that could facilitate effective implementation and utilization. According to Moody et al, 2004, healthcare practitioners believe that EMR systems could result in a decreased workload. However, in spite of the positive effects of integrated EMR system usage in medical practices, the adoption rate is still low and often meets resistance from many stakeholders in healthcare especially clinical staff. The attitude of healthcare towards integrated EMR systems would depend on their prior exposure to such systems and the systems' characteristics such as ease with which the use of the system can learned, its user friendliness and availability of user support (Albert & Manda, 2010).

Cost and availability of funds for the acquisition of the required technological infrastructure could be another influential factor for the diffusion of integrated EMR systems (Khalifa, 2013). Various health institutions have made substantial efforts toward adoption and meaningful use of EMR system but are faced with financial challenges such as hardware, software, implementation, broadband internet connectivity and other equipment costs (Rihab & Helen, 2014). The dwindling government funding for such projects present significant challenges for effective diffusion on integrated electronic medical records systems.

Establishment and enforcement of policy, legal and regulatory framework are also critical to the success of integrated EMR systems. These policies, standards and regulatory frameworks would guide and ensure maintenance the privacy, security and confidentiality of patient data among all users, which are of utmost importance to both patience and clinicians (Albert & Manda, 2010). Various security measure,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

standards and their compliance at every level of usage is paramount. This require security and systems training for all stakeholders with access to such integrated electronic medical records (HealthIT.gov, 2014).

Challenges of Integrating Electronic Medical Records Systems

Although there are significant evidence of individual healthcare facilities in Ghana adopting one form of electronic medical records or the other, they remain siloed in these institutions with low level of integration among departments in the same health institutions, among hospitals, clinics and among other institutions playing vital roles in the health delivery system such as laboratories and pharmacies (Acquah-Swanzy, 2015).

According to Kristiina et al., 2008, one of main factors militating against the integrating EMR systems is the issue of interoperability: Many different EMR projects do not take into account the need to ensure interoperability with other similar systems. This is often due to the absence of standardization of technologies and implementation platforms such as database management systems and operating systems to be used. A further challenge is the use different in terminologies making interoperability difficult (Ashish et al., 2009)

There also the problem of health institutions using different systems and software, whilst some are using desktop application, others may be using web-based applications or mobile applications. This makes information and health record sharing nearly impossible.

The low levels of adoption of electronic health records in health institutions suggest that policymakers face substantial obstacles to the achievement of healthcare performance goals that depend on health information technology (Faustine & Suzanne, 2008). A policy strategy focused on financial support, interoperability, and training of technical support staff may be necessary to spur the integration of EMR systems (Khalifa, 2013).

Policy Frameworks for Integrated Electronic Medical Records System in Ghana

In Ghanaian, the ministry of health has developed a number of policy frameworks to facility the integration of ICT into the health delivery systems (see Figure 1). Prominent among these policies is the five year programme (2002 to 2006) which had the objectives to improve overall health status and reduce inequalities in the health outcome of people living in Ghana through the use of ICT (Afagbedzi, Obuobi, Aryeetey, & Bosomprah, 2013). The government of Ghana's information and communication technology policy document for accelerated development (ICT4AD) states among others, that, ICT would be utilized to promote national health. The national health ICT policy thus provides the strategic roadmap to deploy an integrated EMR systems to improve the quality of healthcare outcomes. The policies aimed at improving the information and communication technology infrastructure in the health sector by networking all health institutions and providing adequate ICT tools for service delivery and management (MOH, 2005).

The Conceptual Framework for the Study

A conceptual model (Figure 2) based on Technology- Organizational- Environmental framework (Tornatzky & Fleischer, 1990) for innovative technology diffusion is proposed for the study. Using the available literature on the electronic medical records diffusion and integration in healthcare institutions,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Healthcare ecosystem in Ghana

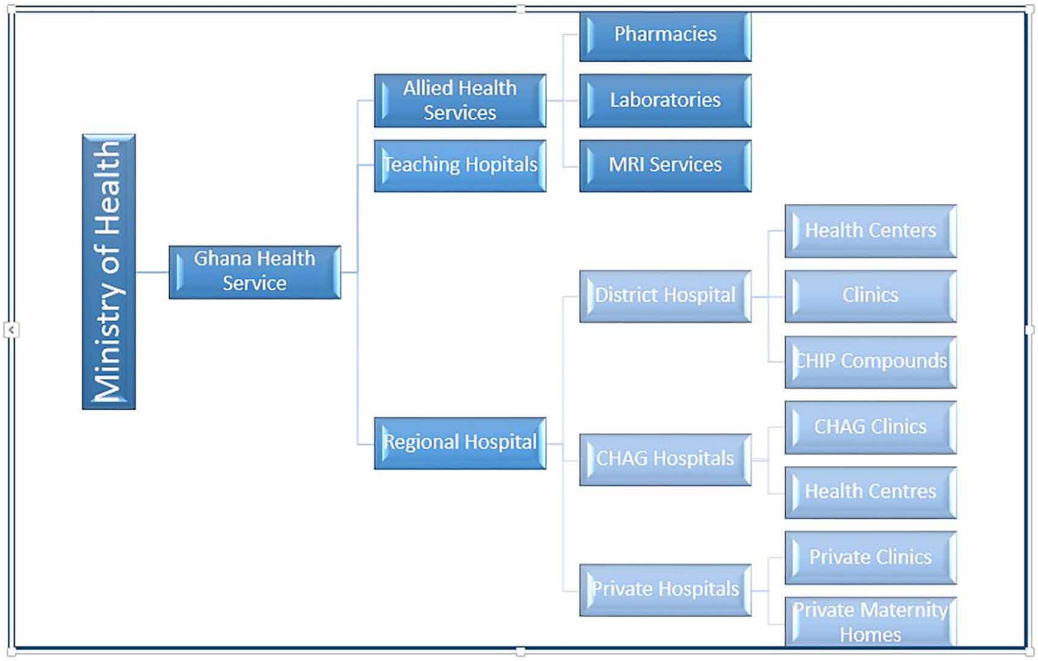
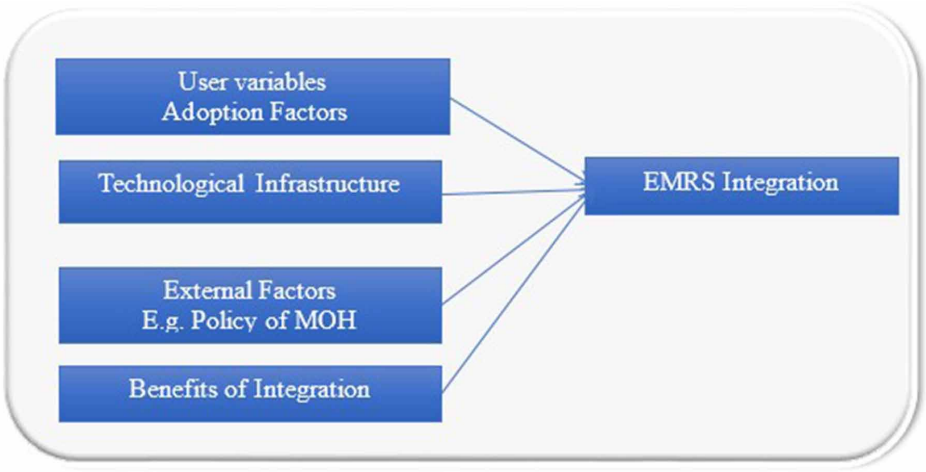


Figure 2. Conceptual model adapted from (Tornatzky & Fleischer, 1990) (Agarwal & Prasad, 1997)



the conceptual framework identified and categorized various factors that could potentially influence the integration of electronic medical records in Ghana.

The study thus, categorizes these main factors into four criterion variables namely, User variables, Existence of Technological Infrastructure, External Factors and Integration Benefits. There are evidences from various studies indication a correlation between these variables and effective integration of electronic medical records for effective healthcare delivery.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

User Variables

These are the various factors relating to healthcare professionals and their demographic characteristics. The variables include the level of the end users' practical knowledge in computer usage which is measure by their level of skill in computing. This categorized them into basic, intermediate, advance and proficiency levels in computing skills (Abell, Bragg-Underwood, Alexander, Abell, & Burd, 2015):

1. The Basic level denotes users are able to use the computer for typing and use of office applications;
2. Intermediate level denotes users who can make use of the internet for sending e-mails and research activities as well as share files;
3. Advanced level denotes user who are able to use or quickly adapt to the use of both web-based and desktop database applications;
4. Proficiency level: Users at this level are able to combine all the skills above their daily activities (Beiter, Sorscher, Henderson, & Talen, 2008).

Technological Infrastructure

Technological infrastructure combines the both hardware and software required for the effective deployment and utilization of integrated ERM systems. These include includes client computer systems, printers, backup drives, servers, uninterrupted power supply (UPS). Availability of network resources which includes the Internet access, switches. Availability of software systems which includes the server operating system, client operating system, application software and database management systems (Hoque, Rakibul, & Golam, 2016).

External Factors

These comprise laws, policies and regulations of the Ministry of Health to support and facilitate the integration of EMR systems (Gilad, 2011). The government of Ghana's information and communication technology policy document for Accelerated Development has stated that, ICT should be utilized to promote national health in order to formulate strategic roadmap to achieve the vision and goal of national health ICT policy. These policies improve the information and communication technology infrastructure in the health sector by networking all health institutions and by providing adequate ICT tools for service delivery and management ((MOH, 2005).

Perceived Benefits of Integrated ERM System

Integrated Electronic Medical Record (EMR) systems are becoming an essential part of the fabric of modern healthcare. EMR systems have evolved from pure record-keeping to an integrated, enterprise-wide system that holds the promise of accurate, real time access to patient healthcare data (Wiliam & Michael, 2009). It is believed stakeholders appreciation of the potential benefits of integration would also facilitate acceptance of the integration (Wang, et al., 2003).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This assesses the stakeholders' perceptions on the important role of integrated EMR systems in the provision of quality healthcare such as data quality, information sharing among healthcare institutions, easy retrieval of patient information in order to ensure continuity of care, avoidance of prescription errors and cost savings in healthcare provisioning (Nguyen, Bellucci, & Nguyen, 2014).

METHODOLOGY

The study adopted exploratory cross-sectional study designed. Cross sectional study make inference by basing on the overall findings to generalize for the entire population (Creswell, 2009). Purposive sampling was used to collect data from the targeted 200 health information officers and managers of healthcare institutions and managers across the country through both paper-based questionnaire and email. Quantitative method was used with the variables assigned numeric values on a Likert scale (Neuman, 2007). Permission was sought from the various hospital administrators before the questionnaires were administered. Respondents were assured that all information provided were for academic purpose and would be handled confidentially.

DATA ANALYSIS AND PRESENTATION

The data analyses was carried out with Statistical Package for Social Sciences (SPSS version 21) using both descriptive and inferential statistics. Factor analysis was done to find out the influential factors influencing the integration of EMR Systems in Ghana. Out of the 200 questionnaire administered, 150 responses were received and analysed quantitatively using descriptive and inferential statistics.

Demographics of the Respondents

Table 1 presents the demographic characteristics of the respondents of the research. Out of 150 respondents, 51.3% were between the ages of 18-29 years, 40.7% between 30-39 years, and 8% were 40 years and above.

The gender distribution of respondents shows 52.7% males and 47.3% females.

With respect to educational background of the respondents, 53.3% had the Certificate, Diploma and HND qualification, 38.0% Degree, 4.7% SHS and A&O Level, 2.0% basic education, 2.0% also indicated others.

In terms of positions, 32.7% were Records and Health information officers, 21.3% Physicians and Physician Assistants, 25.0% were managers and administrators whilst Laboratory technicians made up 10.7% and 10.3% were pharmacist and pharmacy assistants.

Computer Usage Proficiency

To assess the user variables, the level of computer literacy among the respondents and their exposure electronic medical records in their institutions were assessed. The responses and the analyses are presented in Table 2.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Demographics of the respondents

Variable	Frequency	Percent
18-29	77	51.3
30-39	61	40.7
40 and above	12	8.0
Total	150	100.0
Variable	**Frequency**	**Percent**
Male	79	52.7
Female	71	47.3
Variable	**Frequency**	**Percent**
Basic Level Education	3	2.0
Secondary/ High School Education	7	4.7
Diploma and HND	80	53.3
Degree	57	38.0
Post Graduate	3	2.0
Variable	**Frequency**	**Percent**
Records and Health information officers	49	32.7
Physicians/Assistants	32	21.3
Managers/Administrators	38	25.3
Lab technicians	16	10.7
Pharmacist/Assistants	15	10.0
Total	150	100.0

Table 2. Level of computer literacy

Variable	Frequency	Percent
Basic	20	13.3
Intermediate	27	18.0
Advanced	22	14.7
Proficiency	81	54.0
Total	150	100.0

Out of the 150 respondents, 20, making 13.3% have basic computer knowledge such working with office applications, 27, 18% have intermediate knowledge such as using emails, 22, 14.7% have advanced knowledge including manipulating data in excel and a greater majority, 81 making up 54.0% are proficient computer users indicating their work schedule requires the use of computers.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Existence of EMR Systems in the Facilities

The Tables 3 and 4 presents the means and the type of data collected and how data is shared in the respondents' facilities. The extent of the use of computers in data collection and sharing data electronically would indicate that the respondents are exposed to the operations of EMR Systems. This would facilitate easy diffusion and acceptance of intergraded ERM system if it is made available.

The responses analysed in Table 3 shows that majority representing 48% capture data both manually and electronically, electronic only was 38.0%, and manual only was 14%. This is an indication that majority of the health care facilities are conversant with electronic medical recording systems. However, 14% still using manual systems in on the high side.

Table 4 presents the analysis of the extent of data sharing practices of the health institutions surveyed. Manual data sharing had a mean score of 3.09 with a standard deviation of 1.400. Data shared electronically has mean scores of 3.24 with a standard deviation of 1.432. Data stored centrally on servers had a mean score of 3.51 with 1.24 standard deviation. Whereas easy access to patients' data had a mean score of 3.99 with standard deviation of 1.033. The generally high standard deviations mean that the values in the responses from the respondents are farther away from the mean score (Rumsey, 2016).

Perception of the Potential Benefits of Electronic Records System

Tables 5 presents respondents' perceptions of the potential benefits of integrated electronic medical records, the study sort to find out the extent to which electronic medical records system could benefit their healthcare facilities and improve the quality of healthcare delivery.

First, perception of the respondent on the effect of integrated EMR systems on quality of healthcare data was ascertained using the characteristics data quality (Raghupathi & Raghupathi, 2014).

Table 3. Means of data capture in health institutions

Variable	Frequency	Percent
Only manual	21	14
Only electronic	57	38.0
Both	72	48.0
Total	150	100.0

Table 4. Data sharing within the facilities

Variable	N	Mean	Std. Deviation
Data Manually Shared	150	3.09	1.400
Data Electronically Shared	150	3.24	1.432
Data Stored Centrally on Servers	150	3.51	1.241
Data Stored on Individual Department	150	3.99	1.033
Valid N (listwise)	150		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Perceived benefits of electronic medical record systems

Variable	N	Mean	Std. Deviation
Timely data for decision making	150	3.83	.995
Provision of Accurate data	150	3.53	1.001
Generates Consistent Data	150	3.71	1.007
Generates Complete data	150	3.92	.855
Improvement in communication	150	3.70	.857
Facilitates workflow	150	3.79	.966
Ensures patient safety	150	3.79	1.019
Patient access to data	150	3.09	1.300
Patient waiting time has reduced	150	3.26	1.178
Improve diagnosis and outcome	150	3.69	.941

Table 5 presents the analysis of the responses on perception of the respondents of the effect of integrated EMR systems on data quality for the healthcare delivery. Provision of timely data for decision making had mean score of 3.83 with a standard deviation of 0.995. Generate complete and comprehensive data had a mean score of 3.92 with a standard deviation of 0.855. The provision of consistent data to support health delivery had a mean score of 3.71 with a standard deviation of 1.007. Whilst provision of accurate data to improve diagnoses has a mean score of 3.53 with a standard deviation of 1.007.

Integrated EMR systems improve communication among staff in healthcare facilities had a mean 3.70 with a standard deviation of 0.857. Integrated EMR systems facilitate workflow had a mean score of 3.79 with a standard deviation of 0.966. Patients' access to their health records is encouraged by the EMRs had a mean score 3.09 with a standard deviation of 1.300. Improve diagnose and outcome of condition with a mean score of 3.69 with standard deviation of 0.941, reduction in patients' waiting times had a mean score 3.29 with 1.178 standard deviation, whilst ensuring patients' safety had a mean score of 3.79 with 1.019 standard deviation. Again, the generally high standard deviations mean that the values in the responses from the respondents are farther away from the mean score.

The Extent Integration of EMR Systems

The study south to establish the extent of inter-connectivity among the healthcare facilities to facilitate effective exchange of patients' information. The analysis of the result is presented in Table 6.

From Table 6, the Facilities connecting other facilities had a mean score of 2.53 with a standard deviation of 1.35. Sharing health records facilitate exchange information with other health care facilities had a mean score of 2.52 with a standard deviation of 1.422. Sharing the same EMRS with other facilities had a mean score of 2.57 with standard deviation of 1.051. Able to access the health record of patients' facilities using different system had mean 2.52 with a standard deviation of 1.273. Connecting to other facilities through the internet had a mean score of 2.21 with a standard deviation of 1.244, Connecting through Virtual Private Network had a mean score of 2.44 with a standard deviation of 1.234 whilst Facility used web-based electronic records management had a mean score of 2.91 with a standard deviation of 1.274.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. The extent of integration of EMR systems

Variables	N	Mean	Std. Deviation
EMRS connects to other facilities	150	2.53	1.354
Sharing Health Records with other facilities	150	2.52	1.422
Share EMRS with other Facilities	150	2.57	1.051
Access patients' records on different system	150	2.52	1.273
Connected to other facility via Internet	150	2.21	1.244
Connected to other facility via VPN	150	2.44	1.234
Uses web-based EMRS	150	2.91	1.274
Valid N (listwise)	150		

Factors Affecting Integration of Electronic Medical Record Systems

To determine the factors affecting the integration and utilization of integrated electronic medical record (EMR) system in the health facilities, the respondents were made to rate the various problems identified from previous studies as to the extent to which the factors affect the diffusion of integrated EMR systems. The responses are analysed in Table 7.

Inadequate capital for the purchase of database is agreed since the mean 3.58 with a standard deviation of 1.271. The cost of acquisition and maintenance of EMR is one of the factors influencing the integration of electronic medical records (EMR) systems since the mean 3.57 with a standard deviation of 1.077. Unavailability of staff with adequate expertise in information technology is agreed since the mean 3.25 with a standard deviation of 1.357. Unclear return on investment on EMRs is agreed since the mean 3.31with a standard deviation of 1.216.

Table 7. Factor affecting EMR systems' integration

	N	Mean	Std. Deviation
Funds for Hardware Acquisition	150	3.58	1.271
Applications and Maintenance costs	150	3.57	1.077
Unclear Return on Investment	150	3.31	1.216
Availability of ICT expertise	150	3.25	1.357
Usage of different systems and different platform	150	3.59	1.124
Absence of Standards	150	3.65	1.135
Lack of interoperability with systems in other facilities	150	3.45	1.308
Resistance of physicians and health workers	150	3.35	1.337
Lack of government policies for integrated EMRS	150	3.77	1.172
Security and privacy concerns of patient's information	150	3.97	1.074
Patients' rights on the release of their information	150	3.89	1.267
Lack of Enforcement Data Protection laws	150	4.21	.992
Valid N (listwise)	150		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A Principal Component Analysis of the Influential Factors

Principal components analysis is a method of data reduction. Interpretation of the principal components analysis helps in determining which of the identified variables are most strongly correlated with the determinant variable. The variables with total coefficient farthest from zero in either positive or negative direction in the initial Eigenvalues is considered the most influential factors. Which numbers are considered to be large or small is a subjective decision. There is the need to determine at what level the correlation value will be of importance. In this study, any factor with an initial Eigenvalues above 0.5 is deemed important (Liwicki, Tzimiropoulos, Zafeiriou, & Pantic, 2013).

From Table 8, out of the 12 components, only seven factors were found to have significant influence on the diffusion of Integrated EMR systems in Ghana. These are inadequate capital (4.493), Maintenance of database (2.149); Unclear return on investment (1.338), Unavailability of staff with adequate expertise (1.014), Usage of different systems (.633), absence of standards (.621) and lack of interoperability with initial Eigen value more than 0.5. In factor analysis, if the initial Eigen value is more than 1, it means that, that variable is very significant. This means that, these four components constitute about 75% which gives a clue of the significances of the four factors whiles the rest of the components constitute 25%.

DISCUSSION OF RESULTS, FINDINGS AND CONCLUSION

This study was carried out to analyses the factors influencing the integration of EMRs in the healthcare delivery systems in Ghana. The study identified four key categories of variables that could influence successful diffusion on integrated Electronic Medical Records Systems in Ghana namely User Variables,

Table 8. Principal component analysis of factors

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.493	37.442	37.442	4.493	37.442	37.442	3.185	26.545	26.545
2	2.149	17.909	55.351	2.149	17.909	55.351	2.252	18.768	45.314
3	1.338	11.150	66.501	1.338	11.150	66.501	1.816	15.131	60.445
4	1.014	8.450	74.951	1.014	8.450	74.951	1.741	14.506	74.951
5	.633	5.273	80.223						
6	.621	5.176	85.400						
7	.525	4.379	89.779						
8	.416	3.466	93.244						
9	.310	2.582	95.827						
10	.225	1.872	97.699						
11	.160	1.332	99.031						
12	.116	.969	100.000						

Extraction Method: Principal Component Analysis.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Technological Infrastructure, External factors and Benefits of Integration. These factors were studied from the institutional perspectives.

User Variables

Socio-demographic characteristics obtained from the health institutions surveyed indicate that the health facilities employee ad relatively younger majority of whom possess tertiary academic qualification. This indicate youthful group who are easily adaptable to modern ICT platforms in their work environment. This highly educated workforce would make diffusion of any technological innovation in the healthcare delivery systems potentially successful (Marcelo, da Silva & do Amaral Moretti., 2014).

The level of computer literacy among the staff of the institution surveyed was very high with majority indicating they are proficient users. This would have positive influence on the deployment of Integrated Electronic Medical records systems for Healthcare in Ghana. This confirms (Fustian & Suzanne, 2008) which found the successful adoption of EMRS is influenced by well-educated human resource and youthful computer literate workforce.

Technological Variables: Existence EMR Infrastructure

In spite of the computer proficiency of the staff, data collection and data sharing are predominantly manual operations in the institutions with very minimal electronic data sharing in the institutions among various institutions in the healthcare ecosystem. There is average technological infrastructure that could facilitate integrations of EMR systems in the predominantly urban based healthcare facilities, whereas the rural facilities lack such infrastructure. There is low level interoperability of electronic medical record among departments of the same health institutions and among individual institutions. The electronic health records systems used in various health institutions differ. All these would negatively influence the integration of EMRS. However, the availability of networks and internet connectivity found in various institutions can positively influence the diffusion of integrated EMS.

Perceived Benefits of EMR Systems

The knowledge of the respondents of the various benefits of Integrated Medical records was significantly high. Among the benefits identified are provision of timely data for decision making, provision of accurate and timely data among the benefits. The knowledge of these potential benefits would have positive influence on the acceptance and use of integrated medical records once implemented.

The Institutional and External Variables

Out of the twelve institutional and external factors identified as influencing the integration of EMR, seven of them were statistically influential on the diffusion of integration of Electronic medical records (EMR) systems in Ghana. These are, inadequate capital for acquisition hardware required for diffusion of integrated EMR system, the cost applications and maintenance of EMR systems, unclear return on investment of the integrated EMR systems and unavailability of staff with adequate and requisite expertise to implement and manage EMR systems in the healthcare facilities, lack of standards, lack of interoperability of different systems currently in use in the healthcare institutions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Government policy and support will therefore be crucial in the acquisition and successful diffusion of EMR systems in the healthcare delivery systems in Ghana.

CONCLUSION

The findings of the study points to favorable condition for the diffusion of integrated electronic medical records (EMR) system in the healthcare delivery system in Ghana. However, the absence of enforceable policy by government, the cost of acquisition and maintenance of the systems as well as non-availability of technically skilled manpower in the healthcare institutions are seen as the predominant challenges.

It is therefore recommended that there should be a policy direction for the implementation of integrated electronic medical records in the country to ensure standardization of application, hardware and platforms to be used by all health care delivery. This would ensure interoperability among healthcare institutions in the country. Given the potential benefits identified, government should provide funding for nationwide mandatory diffusion of Integrated EMS throughout the healthcare delivery system in Ghana. This would reduce the burden of cost on the facilities and also ensure availability of technical staff to ensure smooth operation of the systems in Ghana.

REFERENCES

healthit.ahrq.gov. (2015, February). Electronic medical record systems. Retrieved December 05, 2015, from https://healthit.ahrq.gov

Abell, C. H., Bragg-Underwood, T., Alexander, L., Abell, C. E., & Burd, V. (2015). Nurses' Knowledge and Attitudes toward Implementation of Electronic Medical Records. *International Journal of Faith Community Nursing*, *1*(3), 74.

Acquah-Swanzy, M. (2015). Evaluating Electronic Health Record Systems in Ghana: the case of Effia Nkwanta Regional Hospital. Munin Open Research Archive.

Adjorlolo, S., & Ellingsen, G. (2013). Readiness Assessment for Implementation of Electronic Patient Record in Ghana: A Case of University of Ghana Hospital. *Journal of Health Informatics in Developing Countries*.

Afagbedzi, S., Obuobi, H., Aryeetey, R., & Bosomprah, S. (2013). A Review of Ghana's E-health Strategy. In *8th Health Informatics in Africa Conference (HELINA 2013)*. JHIA.

Agarwal, R., & Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences*, *28*(3), 557–582.

Albert, B., & Manda, B. (2010). Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. *BMC Health Services Research*.

Ashish, K. J., DesRoches, C. M., Campbell, E. G., Karen, D., Sowmya, R. R., Ferris, T. G., ... Blumenthal, D. (2009). Use of Electronic Health Records in U.S. Hospitals. *The New England Journal of Medicine, 360(16),* 1628–1638. PMID:19321858

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Beiter, P., Sorscher, J., Henderson, C., & Talen, M. (2008). Do electronic medical record (EMR) demonstrations change attitudes, knowledge, skills or needs? *Informatics in Primary Care, 16(3),* 221–227. PMID:19094409

Boateng, V. (2009). *E-Government and E-Health Strategies*. Accra: ECOSOC Africa Regional Ministerial Meeting on e-Health.

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, California: Sage Publications.

Faustine, W., & Suzanne, A. B. (2008). The role of electronic medical record in care delivery in developing countries. *International Journal of Information Management, 28(6),* 503–507.

Gilad, J. K. (2011). Health-information exchange: Why are we doing it, and what are we doing? *Journal of the American Medical Informatics Association, 18(5),* 678–682. PMID:21676940

Häyrinen, K., Saranto, K., & Nykänen, P. (2008). Definition, structure, content, use and impacts of electronic health records: a review of the research literature. International journal of medical informatics, 77(5), 291-304. PMID:17951106

Healthit.gov. (2014, August 29). Retrieved October 3, 2015, from https://www.healthit.gov/providers-professionals/electronic-medical-records-emr

Hoque, M., Rakibul, Y., & Golam, S. (2016). Investigating factors influencing the adoption of e-Health in developing countries: A patient's perspective. *Informatics for Health & Social Care*. PMID:26865037

Khalifa, M. (2013). Barriers to health information systems and electronic medical records implementation. A field study of Saudi Arabian hospitals. *Procedia Computer Science*, *21*, 335–342. doi:10.1016/j.procs.2013.09.044

Kurtin, S. E., Harrison, L., Iraca, T., & Hassan, A. (2017). 223-Pilot Testing of a Health and Life Management Mobile Health Application for Caregivers and Patients Living with Myelodysplastic Syndromes. *Leukemia Research*, *55*, 129. doi:10.1016/S0145-2126(17)30336-3

Liwicki, S., Tzimiropoulos, G., Zafeiriou, S., & Pantic, M. (2013). Euler principal component analysis. *International Journal of Computer Vision*, *101*(3), 498–518. doi:10.100711263-012-0558-z

Marcelo, G., da Silva, D., & do Amaral Moretti, S. (2014). The Y Generation Myth: Evidences Based on the Causality Relations Among Age, Diffusion and Adoption of Technology of College Students of São Paulo State. *Future Studies Research Journal: Trends and Strategies*, *6*(1), 2–52.

MOH. (2005). *Health Sector ICT Policy and Strategy*. Accra, Ghana: Ministry of Health.

Neuman, W. L. (2007). Basics of Social Research, Qualitative and Quantitative Approaches (2nd ed.). Boston, MA: Pearson Education, Inc.

Nguyen, L., Bellucci, E., & Nguyen, L. T. (2014). Electronic health records implementation: An evaluation of information system impact and contingency factors. *International Journal of Medical Informatics*, *83*(11), 779–796. doi:10.1016/j.ijmedinf.2014.06.011 PMID:25085286

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Price, M. (2013). Adopting electronic medical records. Are they just electronic paper records? *Canadian Family Physician Medecin de Famille Canadien*, *57*(7), e322–e329. PMID:23851560

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, *2*(1), 3.

Rihab, A. H., & Helen, C. (2014). Solutions to overcome technical and social barriers to electronic health records implementation in Saudi public and private hospitals. *Journal of Health Informatics in Developing Countries*.

Rumsey, D. J. (2016). *Statistics For Dummies* (2nd ed.). Wiley.

Tang, P. C., Ash, J. S., Bates, D. W., Overhage, J. M., & Sands, D. Z. (2006). Personal health records: Definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*, *13*(2), 121–126. doi:10.1197/jamia.M2025 PMID:16357345

Tornatzky, L. G., & Fleischer, M. (1990). *The Process of Technological Innovation.* Lexington Books.

Wang, S. J., Middleton, B., Prosser, L. A., Bardon, C. G., Spurr, C. D., Carchidi, P. J., ... Bates, D. W. (2003). A cost-benefit analysis of electronic medical records in primary care. *The American Journal of Medicine*, *114*(5), 397–403. doi:10.1016/S0002-9343(03)00057-3 PMID:12714130

Wiliam, M., & Michael, W. (2009). Integrated Electronic Medical Record Systems: Critical Success Factors for Implementation. In *Proceeding of the 42nd Hawaii International Conference on system sciences*, Hawaii, USA.

This research was previously published in the International Journal of Computers in Clinical Practice (IJCCP), 2(2); edited by Athina A. Lazakidou; pages 38-55; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 7
SOA Governance in Healthcare: Beyond Early Ideas to a Structured Framework

Konstantinos Koumaditis
University of Piraeus, Greece, & Aarhus University, Denmark

Marinos Themistocleous
University of Piraeus, Greece & University of Coimbra, Portugal

ABSTRACT

This chapter depicts the gradual development of a conceptual SOA Governance Framework tailored for healthcare organisations. The proposed framework presented herein is based on the authors' previous research and includes nine SOA Governance elements that need to be considered during the SOA process. The identification and conceptualisation of the elements were grounded in the normative literature and further developed to include healthcare specific aspects. This comes as a method to overcome the limitations identified in normative literature and enhance the elements' conceptualisation. Besides, the authors propose a unique design combining nine elements of SOA Governance with SOA Critical Success Factors (CSFs) and Healthcare Information Systems (HIS) challenges. This proposal aims to pinpoint attributes and guidelines for each element, required to successfully govern SOA and tackle longstanding HIS challenges. The framework is intended to be used as a decision supporting tool for SOA Governance in a healthcare setting.

INTRODUCTION

Economic conditions demand that Healthcare Information Technologies (HIS) deliver ever more business value. Today, HIS play a more and more important role in the provision of healthcare. The scope and complexity of HIS projects has increased, requiring excellent project management to fully realize the benefits of such an investment. Unfortunately, healthcare projects are associated with considerable failure rates or having significant cost or budget overruns (Kaplan & Harris-Salamone, 2009). Both re-

DOI: 10.4018/978-1-7998-1204-3.ch007

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

search and anecdotal evidence suggests that many HIS projects struggle to meet functionality and quality targets. Research has identified multiple reasons for these challenges in HIS projects, such as: project escalation, poor risk management, failure to manage user expectations, poor software development or project management processes, or inability to learn from past mistakes and successes (Dwivedi et al., 2013; Lunt et al., 2011). Yet, integrated technologies that have been employed by healthcare organisations, to tackle these challenges, have not shown great levels of adoption due to their barriers, like: (a) high costs, (b) resistance to change, (c) organisational issues, (d) high complexity, (e) large scale of change, (f) time consuming implementation and (g) politics.

Consequently, the necessity for integrated HIS in a safer, interoperable and more manageable environment motivated organizations to consider the adoption of paradigms, such as SOA. SOA can be beneficial for a healthcare organisation as it provides interoperability and integration of the legacy HIS, attributes need in modern healthcare organisations.

Nevertheless, despite SOA benefits Heffner (2009) indicates that, 41% of SOA users in the Global 2000 firms believe that: (a) SOA has delivered less benefit than expected, (b) 17% claim they face problems and (c) will not expand SOA use. This reveals that even though SOA is considered a valuable architectural paradigm its application, efficiency and performance are affected by various factors. These statistics indicate that almost half the companies that implement SOA have not figured out how to benefit from their projects. This is attributed to unclear or weak governance planning (Stephens, 2008). In a recent research on the global status of IT governance standards and models (like COBIT, ITIL/ISO20000 etc.) the findings reveal a tendency to adopt such frameworks, but also a lack of a clear "winner" amongst them. For example, amongst 834 business executives, from 21 countries and 10 industries reveal that ISO20000 or ITIL is referred in 28% of them, while COBIT in 12% (ITGI., 2011). Yet, out of the 839 respondents only 10% have been healthcare executives, thus the percentages drop lower regarding their focus on IT governance in healthcare.

In an attempt to study this issue, we focused on SOA CSFs in healthcare, like: (a) alignment of SOA, organization, human and legal aspects, (b) clear goals set from the beginning of the endeavour, (c) complexity introduced to the system, (d) cost, funding and sponsoring the SOA project, (e) SOA culture that can create support and communication between the stakeholders, (f) experience, skills and training of the employees, (g) governance plan to provide compliance and check services concerning capability, security and strategic business alignment, (h) long-term planning to include reusable services that fit future business, (i) adequate measurement of the compliance and performance, (j) maturity identification and progress of the organization in aspects as IT, organization etc., (k) identification of the right candidate projects/pilots, (l) security risks (data confidentiality, access control), (m) detailed roadmap, (n) adequate standards (e.g. XML, WSDl, REST, HL7 etc.), (o) a team with understanding and experience in change management and clear vision of SOA, and (p) testing of the services and impact (Koumaditis, Themistocleous, Mantzana, & Souliotis, 2012).

The outcome of the review, demonstrates that the most frequently reported CSF in the cases reviewed is SOA Governance. In a nutshell, they report that SOA implementations require governance mechanisms to excel, otherwise the architecture will end up complex, uncontrolled, brittle and eventually discarded (Marks, 2008). A failure in a healthcare organisation's IT infrastructure is not an option as the literature is full of cases where healthcare IT failures cost patients' lives (Fitzgerald & Russo, 2005; Johnson, 2011; Kaplan & Harris-Salamone, 2009). This motivates us to study SOA Governance in healthcare.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

SOA Governance is a research area with many issues (e.g. scope, model, associations) uncharted. In an attempt to fill this literature void, in a previous publication we introduced a set of SOA Governance issues that we were asked to revisit and provide a more structured and detailed approach in this publication (Koumaditis, Themistocleous, & Morabito, 2013). In doing so, resulting in the development of a framework that can be used to support decision-making on SOA Governance issues.

The remaining of the chapter is structured as follows Section 2 presents the HIS need for integration, Section 3 the SOA Governance literature review, Section 4 the SOA Governance Framework in healthcare, Section 5 the research methodology, Section 6 the discussion and Section 7 conclusions and future research agenda.

HEALTHCARE INFORMATION SYSTEMS: NEED FOR INTEGRATION

Healthcare delivery is a complex practice which uses different devices and machines for performing tests, monitoring patients and completing tasks, spread across many organisations, surgery rooms, patients' wards, and medical labs. All these devices need specific vendor based interfaces to communicate with computers. These interfaces are implemented based on specific protocols and much often there is a lack of integration amongst them or an increased cost to integrate them (Rhoads, Cooper, Fuchs, Schluter, & Zambuto, 2010).

Independent initiatives like: (a) Integrating the Healthcare Enterprise (IHE) and (b) Meaningful Use provide standards (e.g. HL7, DICOM, etc.) and guidelines (e.g. EHR) to enhance the interoperability and information sharing among healthcare organisations (HealthIT, 2014; IHE, 2012). An enabler of these standards is the SOA paradigm. An example of this is the recognized Healthcare Services Specification Program (HSSP) program, which is a collaborative effort between standards groups HL7 and OMG to address interoperability challenges within the healthcare sector, operating on SOA (HSSP, 2013). Additionally, HSSP's working groups emphasise on the need of an effective SOA Governance model, which they portray as a critical dimension to SOA implementations, as it becomes the guidance, the oversight, the enforcement, and ultimately the compliance mechanism needed for SOA (HL7/OMG, 2010).

Additionally, HIS integration has been associated with various aspects, amongst others: strategic, social, and/or organisational (Finnegan & Hamid, 2009). In this regard, there is a common tendency to address HIS integration by an overall approach, seen as integrated patient centered care (Leventhal, Taliaferro, Wong, Hughes, & Mun, 2012). Integrated patient centered care reflects on integrated HIS (with elements as e-health services) requiring coordination across professionals, facilities, support systems that is continuous over time and between patient visits (Singer et al., 2011). This approach is observed on national healthcare strategies that encourage patient involvement in their healthcare treatment. Moreover, in the USA and Europe, online personal health records that allow patients to manage their health data have emerged (Black et al., 2011).

For example in Finland, this integration trend can be observed in a legislation that allows citizens to access and interact with their own Electronic Healthcare Records (EHRs), ePrescriptions and auditlogs via the Internet (Ruotsalainen, Iivari, & Doupi, 2008). Following similar legislative opportunities worldwide, patients increase their involvement with HIS. This is a growing involvement, seen in parallel with mechanisms for the collection of information (obtained by mobile and other sources) in order to develop an enhanced, complete and integrated view of citizens health status. The latter is reflected in Electronic Health Records (EHRs) and Personal Health Records (PHRs), which are being enriched

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

*and exploited by different actors and stakeholders (i.e. health and care professionals, citizens, nutrition experts, hospitals, etc) in the health ecosystem. Three general PHR models have been proposed (Detmer, Bloomrosen, Raymond, & Tang, 2008): a) the stand-alone model, b) Electronic Health Record (EHR) system, and c) the integrated one, which is an interoperable system providing linkage with a variety of patient information sources such as EHRs, home diagnostics, insurance claims etc. The main types of health information supported by PHRs are problem lists, procedures, major illnesses, provider lists, allergy data, home-monitored data, family history, social history and lifestyle, immunizations, medications and laboratory tests (Halamka, Mandl, & Tang, 2008; Tang, Ash, Bates, Overhage, & Sands, 2006). Widely known PHR platforms in terms of centralized web-based portals include Dossia (*www.dossia.org*) and Microsoft Health Vault (*www.healthvault.com/*) platforms. Many systems presented in literature offer integration with already established PHRs platforms (Reti, Feldman, & Safran, 2009; Zhou, Yang, Álamo, Wong, & Chang, 2010). Early experiences from the adoption of PHR-based systems have been found to be positive, showing that such systems can be feasible, secure, and well accepted by patients (Jennett & Watanabe, 2006). Nonetheless, today's EHRs and PHRs are far from being what the citizens consider as of value to their health, since for the public view, health means more than being disease-free.*

Following this trend for patients' empowerment, academics, practitioners and patients advocate in favor of the patient centered healthcare systems. Still the aforementioned advocates have not yet reached a concise definition of Patient-Centered e-Health (PCEH) that is shared across the research disciplines that focus on health and Information Technology (IT) (Wilson & Strong, 2014). The lack of consensus can be attributed, amongst other, (a) on the number of challenges that are involved in transitioning healthcare delivery to a more patient-centered system and (b) the lack of proof-of-concept through well-documented and effective PCEH projects.

Many of these applications are based on SOA as HIS can be easily delivered to both desktop and mobile computer devices using for instance Java and Visual Studio development environments. Mobile technologies and especially smart phones support patients and professionals to interact with integrated EHR and other e-health applications. This is an emerging area of e-health, given the term m-health (Chatterjee, Chakraborty, Sarker, Sarker, & Lau, 2009). Sarasohn-Kan (2010) report that a major mobile application vendor provides 5,805 health, medical and fitness applications with 73% of them used by patients and 27% by healthcare professionals (Sarasohn-Kahn, 2010). The aforementioned approaches empower the patients and allows them to take their own measurements, and provide verbal and written inputs (Clemensen, Rasmussen, Denning, & Craggs, 2011).

Clearly, except the introduction of standards (e.g. HL7) all these endevours require adequate healthcare specific SOA Governance to be developed and run in a manageable way. SOA can facilitate the plan to redesign HIS towards patient-centred systems in parallel to end-to-end integration (Bhandari & Snowdon, 2011; Leventhal et al., 2012; Papakonstantinou, Malamateniou, & Vassilacopoulos, 2010). Moreover, Business Process Management (BPM) techniques and tools bridges SOA and aligns business processes (based on patient needs) to end-to-end integration (Koufi, Malamateniou, Papakonstantinou, & Vassilacopoulos, 2010; Strnadl, 2007). Thus, SOA can transform heterogeneous and incompatible systems in a patient-centred environment of cooperating services based on healthcare business processes and services (Bhandari & Snowdon, 2011). SOA can be directly associated with patient centred design, as sub-categories of SOA like User-Centred SOA (UCSOA) can facilitate the redesign of HIS towards a patient centred design including support for the end user (e.g. patient) (Chang, He, Tsai, Xiao, & Chen, 2006; Lizcano, Soriano, Reyes, & Hierro, 2009). Furthermore, SOA can assist in the m-health endeavors by providing interoperable services, support diverse processing efficiently, enable cross-platform

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

communication and adapt dynamically to meet changing needs (Firat, Moser, & Melliar-Smith, 2008; Themistocleous, Azab, Kamal, Ali, & Morabito, 2012). In addition to this, SOA interoperability can be seen as a way to minimise medical errors, by providing data exchange accurately, effectively, securely and consistently (Mantzana, Koumaditis, & Themistocleous, 2011).

The value of SOA can be realised from a management point of view, when SOA incorporates a well-developed SOA Governance framework (Biske, 2008). This way, SOA Governance ensures an organisation is (a) pursuing an appropriate SOA strategy aligned with business goals, and (b) is executing that strategy in accordance with guidelines and constraints defined by a body of SOA principles and policies (Marks, 2008). By focusing on SOA Governance healthcare organisations can better address issues like complexity, security, standards, education/training and even provide the backbone for more capabilities like cloud computing. In the following section the authors address SOA Governance.

SOA GOVERNANCE LITERATURE REVIEW

Literature indicates that literature reviews especially when addressing a new area of research (e.g. SOA Governance) is required to be rigorous and comprehensibly described (Levy & Ellis, 2006). Only then can readers assess the exhaustiveness of the review and other scholars in the field can more confidently (re)use the results in their own research (vom Brocke et al., 2009). Following this view we proceeded to a rigorous review of the normative literature (Koumaditis & Themistocleous, 2013) which resulted in the creation of a SOA Governance concept map, as seen in Figure 1. Due to limited space, the authors will not proceed to analyse the concept map for this publication, but in the following paragraph briefly discuss its use.

The SOA Governance concept map provides not only a graphical representation of the main concepts, but aids in the discussion of the similarities (common elements), differences (scale, detail and analysis) and limitations (confusion, overlapping, and lack of testing and healthcare aspects).

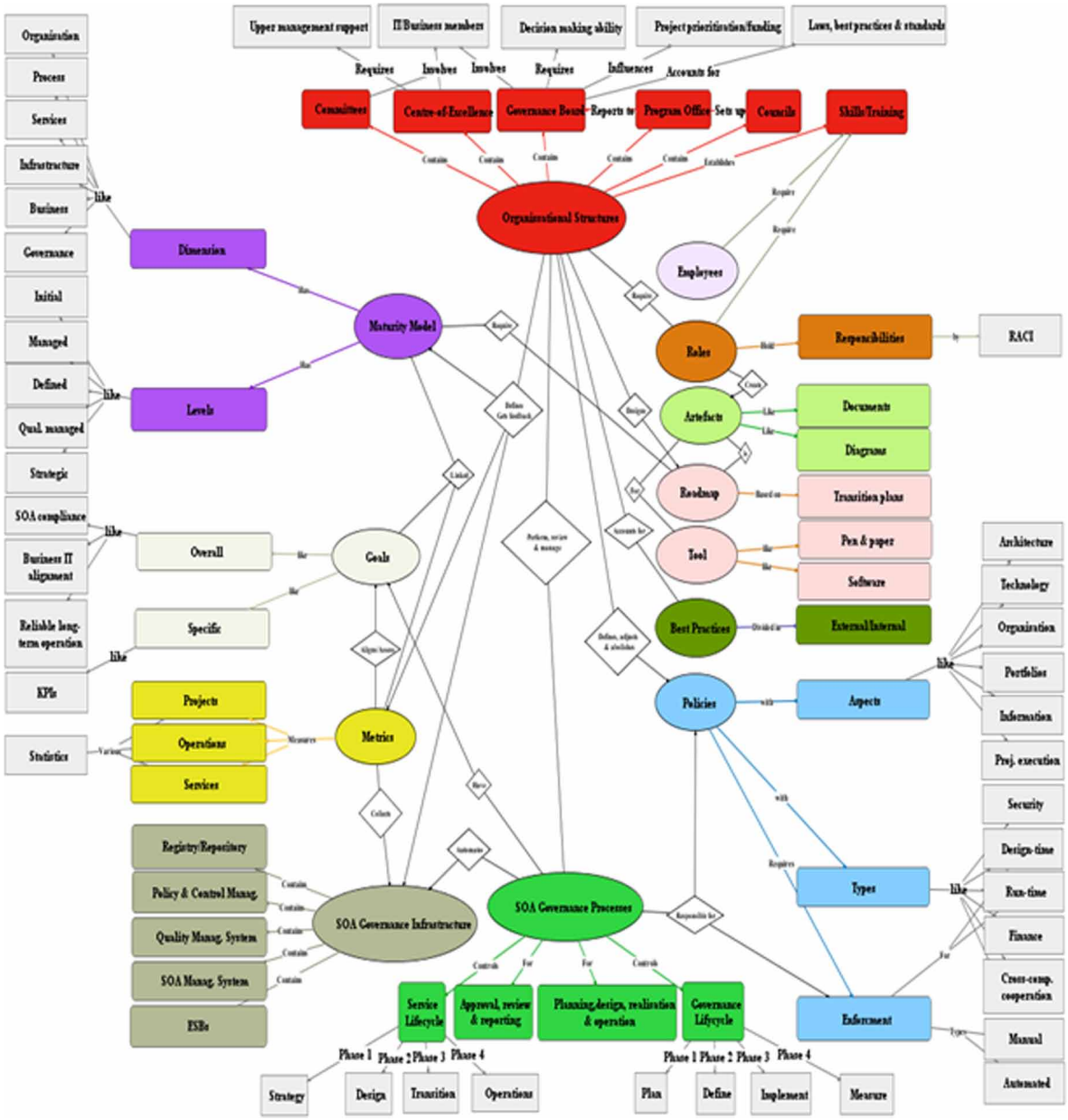
Moreover, the similarities contribute in the construction of the SOA Governance Framework as they provide the common elements and understanding to create the proposition. On the other hand, the differences and limitations as overlapping and confusion, allowed the authors to discard elements as ‘employees’, ‘artefacts’, ‘tool’ and ‘goals’. The decision for exclusion was based on an individual basis. For example, the employees concept overlaps with the organisational structure concept as they both refer to the same notion, while the latter was more frequently observed in the literature and portrayed in more comprehensive conceptualization, thus the first was discarded. The remaining nine common SOA Governance concepts, structure the main elements of the SOA Governance Framework. These elements are:

- **Organisational Structure:** Includes a range of high level structures (e.g. Centre-of-Excellence, boards, offices) and committees/councils dedicated to SOA Governance,
- **Roles:** Concept is identified as newly created or already applied roles with SOA related responsibilities,
- **Maturity Model:** Is an assessment mechanism who provides a comparison between current and future states of SOA aspects (like IT, process, organization),
- **Roadmap:** Concept is portrayed as a transition plan that provides guidance through the SOA endeavor,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **SOA Governance Infrastructure:** Manages governance information, automates processes, enforces policies and collects metrics,
- **Policies:** Comprise rules for deploying and managing SOA with the aid of enforcement mechanisms (manual or automated)
- **Metrics:** Provides visibility of essential parameters of the SOA initiative,
- **Best Practices***:* To collect experiences and supports the successful implementation of SOA,
- **SOA Governance Processes:** Deal with the management and execution of SOA based on the service lifecycle.

Figure 1. SOA Governance Concept Map



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yet, the authors identified limitations in the normative literature as limited scope and analysis of the concepts (e.g. above mentioned nine elements). To this, the authors propose the following additions:

- The *Organisational Structure* needs to be designed as to compensate for the weaknesses of the current organisational IT structures and incorporate a defined engagement approach with the rest of the organization,
- A list of *Roles* needs to be incorporated in this element exemplifying the corresponding responsibilities,
- Design guidelines to aid in the conceptualization of a *Roadmap* are added to this element,
- The Goal/Questions/Metric method (see (Basili et al., 2010) is utilized as a guideline for *Metrics* element,
- As part of the *Best Practices* element a categorization is included,
- *SOA Governance Processes* are portrayed in four phases: plan, design, implement, measure, and additional attributes are provided for each.

Another limitation identified by the authors is that the nine concepts are tailored for a non-specific sector, whereas the authors emphasise on the need to include healthcare specific aspects as a failure in an IT project directly impacts and endangers patients' quality of care and well-being (Fitzgerald & Russo, 2005; Johnson, 2011; Kaplan & Harris-Salamone, 2009). Thus, the authors propose the use of the SOA CSFs as additions to the nine common issues. This unique grouping is presented in the following section of this chapter.

SOA GOVERNANCE FRAMEWORK FOR HEALTHCARE

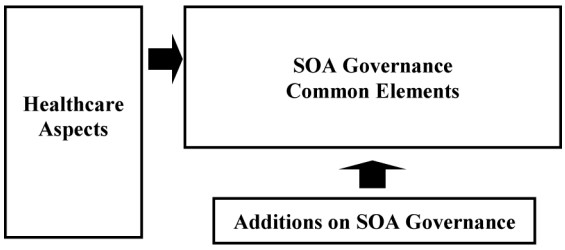
Following the literature review findings of the previous section, the authors proposed the main nine elements that a SOA Governance Framework may consist, namely: (a) Organisational Structures, (b) Roles, (c) Roadmap, (d) Best Practices, (e) Policies, (f) SOA Governance Processes, (g) SOA Governance Infrastructure, (h) Metrics, and (i) Maturity Model. This is identified as the first component of our proposition. The second component are specific additions that were introduced as guidelines to enhance them, like: (a) SOA Governance structure guidelines, (b) example list of SOA roles and responsibilities, (c) SOA roadmap design guidelines, (d) method to construct metrics, (e) best practices categorisation guidelines, and (f) SOA Governance phases and their attributes (examples of the latter). In this section and as a third component, the authors propose specific additions (based on SOA CSFs in healthcare) to the nine elements so as to tackle their lack of healthcare specific conceptualisation. The aforementioned synthesis can be graphically observed in Figure 2.

In more detail, the process of addressing the SOA CSFs is presented in the following paragraphs.

- **Alignment:** Beimborn et al., (2009) suggest that well established SOA Governance structures positively influence the structural dimension of IT/business alignment. The authors agree with the aforementioned view and believe that organisational structures can address this CSF with attributes as dedicated committees and/or individuals who hold alignment related roles and responsibilities (e.g. strategy manager). These individuals can then set specific goals like SOA compliance

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. SOA Governance Framework Synthesis



(e.g. alignment with legal, internal regulations) and business/IT alignment (these goals are also supported by Niemann (2010)), implement SOA Governance processes, design policies that influence the alignment and using appropriate metrics measure and control the level of complexity.

- **Clear Goals:** This is a broadly defined CSF that can be implied for all the elements of the proposed SOA Governance Framework. For example, SOA Governance processes and metrics have specific goals that are set by organisational structures and individuals holding roles in these structures. It is our understanding that the clear definition of these goals is placed on the clarity of the concept they address and the skills of the individuals defining them. Thus, we propose to include in the organisational structure element the guidelines: (a) responsible to set clear goals for all elements and (b) involve individuals with skills that can set clear goals so as to address this CSF.
- **Complexity:** We argue that identifying sources that increase the overall complexity in a system can be placed as a responsibility in individual roles. For example, the SOA architect holds the responsibility for designing and implementing the logical/physical architecture and thus can foresee and manage in some extent the complexity introduced by the implementation. Additionally, individuals with these kinds of responsibilities (e.g. complexity related) can collectively under an organisational structure (e.g. committee, interest groups) or more informally (e.g. intranets, wikis) address potential complexity issues. Yet, as complexity was also identified as a HIS challenge we emphasise the need for this CSF to be adequately addressed. Literature suggestions recommend the utilisation of SOA metric systems that can measure design-time and run-time services and predict unwanted levels of complexity (Hirzalla, Cleland-Huang, & Arsanjani, 2009). We agree with this suggestion and propose that the metrics element (metrics that measure system complexity) need to be utilised so as to provide better governance of the complexity factor.
- **Cost:** Cost and more precisely funding was included in the organisational structure element that can control in certain extent the funding of SOA projects. However, it is our belief that for this CSF to be better dealt the creation of committees in charge of cost planning and control that can also identify service/project sponsors within or outside the organisation, is required. Additionally, funding can be added as an area in the SOA maturity model, so as the healthcare organisation can compare the progress of business processes, technology and other areas with financial data and better govern this factor.
- **Culture:** We argue that the organisational structure can cultivate a SOA friendly culture and gain the trust and cooperation with the rest of the healthcare organisation. This may involve specific governance processes that promote SOA and its beneficial parameters (e.g. educational and communication planning). Thus, we propose the aforementioned additions in the corresponding elements.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Experience and Training:** This CSF including the requirements of experienced personnel and additional training is in great extent covered by the SOA organisational structures' and SOA Governance processes' elements.
- **Governance:** It is our understanding that the proposed SOA Governance elements (Part A and B) and healthcare aspect added in this section (Part C) can address all the requirements for governance (as an CSF) in a healthcare organisation.
- **Long-Term Planning:** We argue that the CSF identified as long-term planning can be represented by the SOA maturity model as a more elaborate approach. For example, one of the application areas that a SOA maturity model examines is the services, these are described starting from an initial and reaching to a final maturity level, expressed in a similar progressive way as in a long-term plan. Thus, we identify this approach to be a useful tool for SOA Governance and propose its utilisation to address this CSF.
- **Measurement:** It can be stated that this CSF can be addressed by the metrics element, as they hold similar conceptualisation, with emphasis on the role of the SOA Governance infrastructure in the collection of metrics and diffusion of feedback about them.
- **Maturity Identification:** The authors accept that this CSF can be represented by the SOA maturity model element.
- **Project Identification:** It can be argued that this factor is addressed as part of the SOA organisational structure and also it is the authors' belief that specific roles, like strategy manager, SOA architect and SOA designer, are responsible for the project identification.
- **Resources:** This CSF is described as a multifaceted one with parameters as financial (budget), technical (software/hardware), and staff (experts, implementation team, consultants) aspects. The financial parameter can be addressed similarly the cost CSF (e.g. from organisational structures and SOA maturity model elements). As far as the technical aspect we believe that is covered by the SOA Governance infrastructure element conceptualisation, since both refer to adequate resources as software and hardware systems. Additionally, the introduction of the roles element can resemble the staffing aspect of this CSF.
- **Risk:** It can be stated that this CSF is addressed by the proposed roles of security manager and tester regarding the security and error risk, and from the organisational structure element regarding the funding and sponsoring risk (in the form of dedicated committees or groups). Yet, it is our understanding that data security and patients' confidentiality are clearly a significant challenge to healthcare and beyond the organisational structures and roles, specific testing, standards, policies and best practices are needed. These will be discussed in the following paragraphs.
- **Roadmap:** It can be stated that this CSF can be addressed by the roadmap element,
- **Roles:** We argue that this CSF can be addressed by the roles element.
- **Standards:** SOA is based on established global standards (e.g. XML, UDDI, SOAP, WSDL, BPEL and REST). In healthcare the standards related to information coding and sharing (e.g. SNOMED, HL7) or protocols (e.g. DICOM) that enhance the interoperability and improve the way HIS share information are usually developed from national central standardisation bodies or independent organisations like IHE (IHE, 2012). This combination between SOA and healthcare standards is an important factor for SOA implementations, regarding aspects of security, interoperability and QoS. Thus, it is our belief that such standards need to be incorporated as best practices and utilised accordingly. In charge of such best practices can be organisational structures (e.g. committees, interest groups) that can identify the standards, evaluate them and recommend

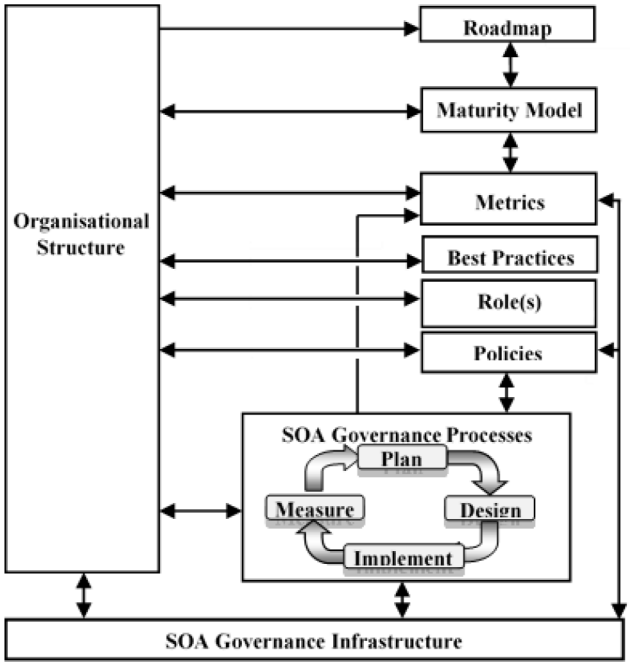
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

them as best practices and thus minimise the risk in issues like: security, access control and medical errors. Additionally, this activity can be included as a SOA Governance process (e.g. standardisation process).

- **Team:** We argue that this CSF can be addressed by the SOA Governance organisational structures element.
- **Testing:** This CSF was identified requiring tools and methods to test the integration approaches and services before the actual service implementation. The authors extents this CSF's importance to testing in both design- and run-time SOA and believes that specific roles like tester, infrastructure and SOA architect, service developer, process and security manager with the aid of SOA Governance infrastructure can control the criticality of this factor. These roles include a wide spectrum of responsibilities related to testing like error localization for service developer, and/or maintenance for process manager and of course testing for the tester role. Yet, it is our understanding that since testing was identified as a HIS challenge, additional emphasis is required to address this factor. Thus, proposes governance processes that address this factor to be included in the SOA Governance processes element, which with the aid of SOA Governance infrastructure can provide testing capabilities. Furthermore, these processes can be placed in the best practices element and form guidelines that minimize errors.

Up till now, most of the SOA Governance elements were introduced in parts, as common elements in Part A, additions to these elements in Part B and healthcare aspects of the elements in Part C. Following is the synthesis of the nine SOA Governance elements, as the proposed SOA Governance Framework, as depicted in Figure 3.

Figure 3. Proposed SOA Governance Framework



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The proposed framework in Figure 3 makes a novel contribution at two levels. Firstly, at a conceptual level, it identifies and proposes nine elements that should be considered during the SOA Governance process. The identification of the nine elements is based on the theoretical works conducted by others in this area (see (Koumaditis & Themistocleous, 2013)). As a result, the conceptualisation of the elements is grounded in the normative literature. In addition to this, the framework incorporates additions that are gathered from additional literature resources beyond the ones that composed the elements. This comes as a method to overcome the limitations identified in normative literature and enhance the elements' conceptualisation. Moreover, the authors propose a unique formulation combining the identified common elements of SOA Governance (seen in Part A and B), the SOA CSFs and HIS challenges. This proposal aims to pinpoint attributes and guidelines for each element, required to successfully govern a SOA and tackle longstanding HIS challenges. Secondly, at a practical level, the proposed framework contributes towards a better understanding of the research area (SOA Governance in healthcare).

RESEARCH METHODOLOGY

To test the SOA governance framework an interpretive qualitative research approach is proposed, with the aid of a case study. The decision to utilize such a research strategy was based on the need to study the SOA governance issues surrounding SOA implementations and on the unique nature of healthcare itself.

In more detail, healthcare can be characterized by unique issues, such as: (a) the interactive nature of practitioner/patient relationships, (b) patients' subjective experiences of illness, (c) their own personal understanding of human suffering and (d) stakeholders behavior and interpersonal interactions (A. Miles, 2009). The aforementioned characteristics of healthcare drive effective clinical process to rely not to a single but rather several different kinds of knowledge (e.g. medical, anthropological and instrumental). Consequently, giving the required depth of understanding that can be applied to a given problem situation. The same principle, as in clinical practice, is useful to be applied to the research methodology so as to research a healthcare setting (Fossey, Harvey, McDermott, & Davidson, 2002). Additionally, as to understand the interplay between stakeholders (clinical and non-clinical) and IS, emphasis is required on the meanings, concepts, definitions, characteristics, descriptions and associations rather than numbers and statistics. Thus, as a more suitable candidate the interpretive qualitative research approach is followed in this research, as it can relate closely to the healthcare aspect of the research. Additionally, qualitative research methods are being used increasingly in evaluations of computer systems and IT both within and outside healthcare (Kaplan & Maxwell, 2005). Furthermore, the use to consider a case study strategy is supported, as it is a widely used research mechanism for examining IS frameworks and supports the investigation of: (a) little-known phenomena (SOA governance in healthcare) and (b) complex processes (SOA implementation) in their natural setting (Bairi & Manohar, 2011).

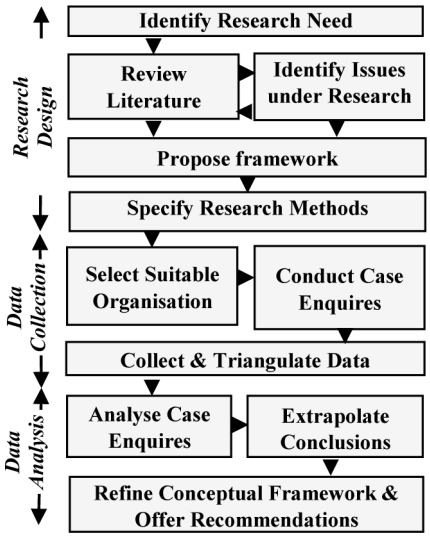
Therefore, the authors propose a qualitative research with three stages: (a) research design, (b) data collection and (c) data analysis, as seen in Figure 4.

DISCUSSION

Healthcare is a unique sector not only due to its relevance to human life, but also due to the big amount of: (a) complexity, (b) legacy systems, (c) data and (d) resources utilized each year. Yet, the develop-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4: Research Methodology Design



ment of SOA Governance in healthcare and its impact to healthcare organizations remains unexposed to research. Such research will undoubtedly provide a more compelling basis for the utilization of SOA and a more controlled environment for HIS integration.

This chapter deals with this important research area and addressed HIS integration challenges. These challenges certainly will be very real over the next few years as governance and management of IT projects, as we continue to embrace SOA projects across widely distributed HIS. To this end and with the help of a rigorous literature review the authors developed and proposed a SOA Governance Framework that can address these challenges. An overview of the proposed SOA Governance Framework was presented in Figure 3. Without a doubt, further research is required to enhance our understanding and impact of such proposition in a healthcare organization. To do this, we already brought forward a qualitative research methodology plan (e.g. Figure 4). Yet, further research actions are needed. These are discussed in the following section.

FUTURE RESEARCH DIRECTIONS

The authors have planned to test the SOA Governance Framework with key SOA players in healthcare such as: (a) public and private hospitals and (b) healthcare organizations (e.g. insurance funds). It is our view that the data gathered from the empirical arena will strengthen, refine and highlight our proposition.

Following the aforementioned actions, we also propose to transform the proposed SOA Governance elements' structure and the framework itself (e.g. refined form) into a large-scale survey questionnaire, instant of using interpretive epistemology. Clearly, this approach would not have been possible previously, since the framework did not exist. Miles and Huberman (1994) discuss such a research design starting with a qualitative approach that provides an initial exploratory phase, whose outcomes may be used to create quantitative instruments (e.g. surveys) to collect new data. It is our understanding that such large-scale survey may offer the opportunity to establish generic significance to the issues related

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to the proposed SOA Governance Framework. In surveying a representative sample of organisations the elements' characteristics related to the proposed framework can be better verified and understood.

REFERENCES

Bairi, J., & Manohar, M. (2011). Critical success factors in gaining user customer satisfaction in outsourced IT services. *Journal of Enterprise Information Management*, *24*(6), 475–493. doi:10.1108/17410391111166530

Basili, V., Lindvall, M., Regardie, M., Seaman, C., Heidrich, J., Münch, J., ... Trendowicz, A. (2010). Linking Software Development and Business Strategy Through Measurement. *Computer*, *43*(4), 57–65. doi:10.1109/MC.2010.108

Beimborn, D., Joachim, N., Schlosser, F., & Streicher, B. (2009). *The Role of IT/Business Alignment for Achieving SOA Business Value - Proposing a Research Model.* Paper presented at the 15th Americas Conference on Information Systems, San Francisco, CA.

Bhandari, G., & Snowdon, A. (2011). Design of a patient-centric, service-oriented health care navigation system for a local health integration network. *Behaviour & Information Technology*, 1–11. doi:10.1080/0144929x.2011.563798

Biske, T. (2008). *SOA Governance*. Birmingham, UK: Packt Publishing Ltd.

Black, A. D., Car, J., Pagliari, C., Anandan, C., Cresswell, K., Bokun, T., ... Sheikh, A. (2011). The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview. *PLoS Medicine*, *8*(1), e1000387. PMID:21267058

Chang, M., He, J., Tsai, W., Xiao, B., & Chen, Y. (2006). *UCSOA: User-Centric Service-Oriented Architecture.* Paper presented at the IEEE International Conference on e-Business Engineering, Shanghai, China.

Chatterjee, S., Chakraborty, S., Sarker, S., Sarker, S., & Lau, F. Y. (2009). Examining the success factors for mobile work in healthcare: A deductive study. *Decision Support Systems*, *46*(3), 620–633. doi:10.1016/j.dss.2008.11.003

Clemensen, J., Rasmussen, J., Denning, A., & Craggs, M. (2011). *Patient Empowerment and New Citizen Roles through Telehealth Technologies - The Early Stage.* Paper presented at the Third International Conference on eHealth, Telemedicine, and Social Medicine, Gosier, France.

Detmer, D., Bloomrosen, M., Raymond, B., & Tang, P. (2008). Integrated personal health records: Transformative tools for consumer-centric care. *BMC Medical Informatics and Decision Making*, *8*(1), 45. doi:10.1186/1472-6947-8-45 PMID:18837999

Dwivedi, Y., Ravichandran, K., Williams, M., Miller, S., Lal, B., Antony, G., & Kartik, M. (2013). IS/IT Project Failures: A Review of the Extant Literature for Deriving a Taxonomy of Failure Factors. In Y. Dwivedi, H. Henriksen, D. Wastell, & R. De' (Eds.), *Grand Successes and Failures in IT. Public and Private Sectors* (Vol. 402, pp. 73–88). Springer Berlin Heidelberg. doi:10.1007/978-3-642-38862-0_5

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Finnegan, D., & Hamid, K. (2009). Information systems (IS) integration approaches in healthcare: a critical review. In W. Currie & D. Finnegan (Eds.), Integrating Healthcare With Information and Communications Technology. London: Radcliffe Publishing.

Firat, K., Moser, L., & Melliar-Smith, M. (2008). Building a Distributed E-Healthcare System Using SOA. *IT Professional*, *10*(2).

Fitzgerald, G., & Russo, N. (2005). The turnaround of the London ambulance service computer-aided despatch system (LASCAD). *European Journal of Information Systems*, *14*(3), 244–257. doi:10.1057/palgrave.ejis.3000541

Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding and evaluating qualitative research. *The Australian and New Zealand Journal of Psychiatry*, *36*(6), 717–732. doi:10.1046/j.1440-1614.2002.01100.x PMID:12406114

HL7/OMG. (2010). The Practical Guide for SOA in Health Care. *Healthcare Services Specification Project.* Retrieved 01.10.10, 1.0, from http://hssp.wikispaces.com/file/view/2008-12-29+SOA%2BHealthcare%2BPractical%2BGuide%2Bv1.0.pdf

Halamka, J., Mandl, K., & Tang, P. (2008). Early Experiences with Personal Health Records. *Journal of the American Medical Informatics Association*, *15*(1), 1–7. doi:10.1197/jamia.M2562 PMID:17947615

Health, I. T. (2014). Meaningful Use Definition & Objectives. *EHR Incentives & Certification.* Retrieved 04.05., 2014, from http://www.healthit.gov/providers-professionals/meaningful-use-definition-objectives

Heffner, R. (2009). *Forrester Report: Survey Results Show SOA Governance Improves SOA Benefit Realization.* Academic Press.

Hirzalla, M., Cleland-Huang, J., & Arsanjani, A. (2009). A Metrics Suite for Evaluating Flexibility and Complexity in Service Oriented Architectures. In G. Feuerlicht & W. Lamersdorf (Eds.), Service-Oriented Computing – ICSOC 2008 Workshops. Springer Berlin Heidelberg.

HSSP. (2013). *Healthcare Services Specification Program*. Retrieved 19.01.2013, from http://hssp.wikispaces.com/

IHE. (2012). *Integrating the Healthcare Enterprise*. Retrieved 01.03.2011, 2011., from http://www.ihe.net/

ITGI. (2011). Global Status Report on the Governance of Enterprise I.T. (GEIT). In Global Status Research. Rolling Meadows, IL: IT Governance Institute.

Jennett, P., & Watanabe, M. (2006). Healthcare and Telemedicine: Ongoing and Evolving Challenges. *Disease Management & Health Outcomes*, *14*(1).

Johnson, C. W. (2011). Identifying common problems in the acquisition and deployment of large-scale, safety–critical, software projects in the US and UK healthcare systems. *Safety Science*, *49*(5), 735–745.

Kaplan, B., & Harris-Salamone, K. D. (2009). Health IT success and failure: Recommendations from literature and an AMIA workshop. *Journal of the American Medical Informatics Association: JAMIA*, *16*(3), 291–299. doi:10.1197/jamia.M2997 PMID:19261935

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kaplan, B., & Maxwell, J. (2005). Qualitative Research Methods for Evaluating Computer Information Systems: Evaluating the Organizational Impact of Healthcare Information Systems. Springer New York.

Koufi, V., Malamateniou, F., Papakonstantinou, D., & Vassilacopoulos, G. (2010). In G. A. Papadopoulos, W. Wojtkowski, G. Wojtkowski, S. Wrycza, & J. Zupancic (Eds.), *Using ESB and BPEL for Evolving Healthcare Systems Towards Pervasive, Grid-Enabled SOA Information Systems Development* (pp. 167–175). Springer, US.

Koumaditis, K., & Themistocleous, M. (2013). *SOA Governance from a Healthcare Perspective.* Paper presented at the IADIS International Conference e-Health 2013, Prague, Czech Republic.

Koumaditis, K., Themistocleous, M., Mantzana, V., & Souliotis, K. (2012). *A Proposition of Critical Sucess Factors Influencing SOA Implementation in Healthcare*. Paper presented at the European Conference on Information Systems, Barcelona, Spain.

Koumaditis, K., Themistocleous, M., & Morabito, V. (2013). New Trends in Healthcare Information Systems (HIS) Integration. In I. M. Miranda & M. M. Cruz-Cunha (Eds.), Handbook of Research on ICTs for Healthcare and Social Services: Developments and Applications (Vol. 1). IGI Global.

Leventhal, T., Taliaferro, P., Wong, K., Hughes, C., & Mun, S. (2012). The Patient-Centered Medical Home and Health Information Technology. *Telemedicine Journal and e-Health*, *18*(2), 145–149. doi:10.1089/tmj.2011.0130 PMID:22304440

Levy, Y., & Ellis, T. J. (2006). A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. *Informing Science Journal*, *9*(9), 181–212.

Lizcano, D., Soriano, J., Reyes, M., & Hierro, J. (2009). A user-centric approach for developing and deploying service front-ends in the future internet of services. *Int. J. Web Grid Serv.*, *5*(2), 155–191. doi:10.1504/IJWGS.2009.027572

Lunt, N., Smith, R., Exworthy, M., Green, S., Horsfal, D., & Mannion, R. (2011). Medical Tourism: Treatments, markets and health system implications: A scoping review. In OECD (Ed.), *International Trade in Health Services* (Vol. 1, p. 55). Paris, France: Directorate for Employment,Labour and Social Affairs.

Mantzana, V., Koumaditis, K., & Themistocleous, M. (2011). *Healthcare IS Interoperability: Challenges and Solutions.* Paper presented at the HEALTHINF 2011: International Conference on Health Informatics, Rome, Italy.

Marks, E. (2008). *Service-Oriented Architecture (SOA) Governance for the Services Driven Enterprise*. Wiley.

Miles, A. (2009). Complexity in medicine and healthcare: People and systems, theory and practice. *Journal of Evaluation in Clinical Practice*, *15*(3), 409–410. doi:10.1111/j.1365-2753.2009.01204.x PMID:19522904

Miles, M., & Huberman, A. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Niemann, M., Miede, A., Wolfgang, J., Repp, N., & Steinmetz, R. (2010). Structuring SOA Governance. *International Journal of IT/Business Alignment and Governance*, *1*(1), 58–75.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Papakonstantinou, D., Malamateniou, F., & Vassilacopoulos, G. (2010). A SOA approach to the implementation of an electronic medical record. *International Journal of Applied Systemic Studies*, *3*(3), 343–352. doi:10.1504/IJASS.2010.034626

Reti, S. R., Feldman, H. J., & Safran, C. (2009). Governance for Personal Health Records. *Journal of the American Medical Informatics Association*, *16*(1), 14–17. doi:10.1197/jamia.M2854 PMID:18952939

Rhoads, J., Cooper, T., Fuchs, K., Schluter, P., & Zambuto, R. P. (2010). Medical device interoperability and the Integrating the Healthcare Enterprise (IHE) initiative. *Biomedical Instrumentation & Technology,* 21-27.

Ruotsalainen, P., Iivari, A., & Doupi, P. (2008). Finland's strategy and implementation of citizens' access to health information. *Studies in Health Technology and Informatics*, *137*, 379–385. PMID:18560100

Sarasohn-Kahn, J. (2010). *How smartphones are changing health care for consumers and providers*. California HealthCare Foundation. Retrieved from http://www.chcf.org/

Singer, S., Burgers, J., Friedberg, M., Rosenthal, M., Leape, L., & Schneider, E. (2011). Defining and measuring integrated patient care: Promoting the next frontier in health care delivery. *Medical Care Research and Review*, *68*(1), 112–127. doi:10.1177/1077558710371485 PMID:20555018

Stephens, E. (2008). *Managing IT investment - The role of EA and SOA with governance*. Paper presented at the SOA in Healthcare realizing quality-of-care, Chicago, IL.

Strnadl, C. (2007). Bridging Architectural Boundaries Design and Implementation of a Semantic BPM and SOA Governance Tool Service-Oriented Computing – ICSOC 2007 (Vol. 4749, pp. 518-529). Springer Berlin / Heidelberg.

Tang, P., Ash, J., Bates, D., Overhage, J., & Sands, D. (2006). Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption. *Journal of the American Medical Informatics Association*, *13*(2), 121–126. doi:10.1197/jamia.M2025 PMID:16357345

Themistocleous, M., Azab, N., Kamal, M., Ali, M., & Morabito, V. (2012). Location-Based Services for Public Policy Making: The Direct and Indirect Way to e-Participation. *Information Systems Management*, *29*(4), 269–283. doi:10.1080/10580530.2012.716743

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). *Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process.* Paper presented at the ECIS 2009, Verona, Italy.

Wilson, V., & Strong, D. (2014). Editors' Introduction to the Special Section on Patient-centered e-Health: Research Opportunities and Challenges. *Communications of the Association for Information Systems*, *34*(15).

Zhou, F., Yang, H., Álamo, J., Wong, J., & Chang, C. (2010). Mobile Personal Health Care System for Patients with Diabetes. In Y. Lee, Z. Z. Bien, M. Mokhtari, J. Kim, M. Park, J. Kim, H. Lee, & I. Khalil (Eds.), Aging Friendly Technology for Health and Independence. Springer Berlin Heidelberg.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

HIS Integration: The alignment of HIS in an interoperable environment.
HIS Interoperability: The ability of HIS to work together and exchange information.
HIS: Information systems designed to facilitate healthcare services.
PCEH: Patient Centered e-health.
SOA Governance: Important mechanisms that ensure the reliable long-term operation of an SOA.
SOA: An architectural paradigm to build ecosystems of services.

This research was previously published in Reshaping Medical Practice and Care with Health Information Systems edited by Ashish Dwivedi; pages 266-286; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 8
KIIPF – An Integrated Inter-Operable Knowledge Management Process Framework for Healthcare: Implementation in Pakistani Healthcare Industry

Mohamad Fauzan Noordin
International Islamic University, Malaysia

Arfan Arshad
International Islamic University, Malaysia & Jazan University, Kingdom of Saudi Arabia

Roslina Othman
International Islamic University, Malaysia

ABSTRACT

Being complex systematically, Knowledge Management (KM) has always been challenging in the healthcare industry (HCI) of developing countries like Pakistan due to gap in acquiring and implementing KM processes. Despite the fact that extensive data and related information is available, the Pakistani HCI has been facing a major challenge due to the absence of a sound and solid KM process framework. Hence, the existing healthcare information systems in Pakistan lack the ability to control and cater the diversified nature of new knowledge creation and innovation. Consequentially, it results in inefficient and ineffective organization of knowledge along with inability to properly utilize important resources that are used for decision making activities. In this study, the researchers present and suggest an integrated and inter-operable KM process framework that helps to streamline and integrate a KM process in a phased comportment that identifies, defines, analyzes, collects and manages knowledge along with its innovation, dissemination, sharing and storage in an appropriate way.

DOI: 10.4018/978-1-7998-1204-3.ch008

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

"A health system consists of all organizations, people and actions whose primary intent is to promote, restore or maintain health" according to the World Health Organization (WHO). Healthcare organizations are composed of health care professionals from multiple disciplines forming several interconnected care teams that strive to provide safe and consistent care. (Ratnapalan & Uleryk, 2014).

The need for an integrated and reliable Knowledge Management process framework in the healthcare industry of Pakistan has become evident due to evolution and innovation of Knowledge Management practices over the past couple of decades. A Knowledge Management Process Framework ensures that all necessary KM elements (Accountabilities, Processes, Technologies and Governance) are in place, and interconnected (Alavi & Leidner, 2001). With a Management Process Framework, KM can take on the aspects of other management systems, and be made part of normal business, rather than relying on an incongruent set of tools (Akhavan, Jafari, & Fathian, 2005). This ensures that there are no gaps in the system, and that knowledge flows freely through the organization.

Health care organizations have to be able to modify their activities based on sudden changes in the condition of their patients or sudden demands due to public health disasters, without compromising patient safety or quality of care. New knowledge creation, technology advances and other market changes can add new and unexpected demands in health care delivery. Health care organizations have to maintain stability following institutional protocols but have to assess their performance and evaluate protocols to create and incorporate new knowledge (Ratnapalan & Uleryk, 2015).

Knowledge management involves the strategies and processes for identifying, capturing, structuring, sharing and applying an individual's or an organization's knowledge to extract competitive advantage and create sources of sustainable growth (Mahmood, Burney, Abbas, & Rizwan, 2012). All the participants of this community create, share and use knowledge to improve the health care service quality and reduce cost. The success of a medical care depends on how effectively and intelligently knowledge is being used to improve the health care process (Mahmood, Burney, Abbas, & Rizwan, 2012).

The importance of Knowledge Management in Healthcare and Medical industry is evident and pertinent (Mahmood, Burney, Abbas, & Rizwan, 2012). Correct and accurate knowledge for the correct and effective treatment of patience is a matter of life and death. The purpose of this research is to propose an outline and present the stages of KM strategy implementation. In addition, it presents the actions necessary at each stage of development of an integrated and interoperable process framework that combines and interrelates KM process steps. This is done while dealing with knowledge effectively and efficiently in any form it exists in the healthcare organization.

BACKGROUND - DEFINING DATA, INFORMATION AND KNOWLEDGE

Knowledge Management (KM) is a term difficult to understand without knowing and understanding the meaning of the word "knowledge". Further, it is also important to clearly differentiate between knowledge, information and data (Botha, Kourie&Snyman, 2008).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Perspective on Knowledge, Information and Data – An Overview

In disciplines like information systems the term "knowledge" is frequently used analogously as information. It is understood as something that can be categorized, classified, organized and communicated as shared knowledge by the help of Information Technology (Jones, 2006). For instance, the encyclopedia at "www.fact-archive.com" defines it as: "information that has a purpose or use."

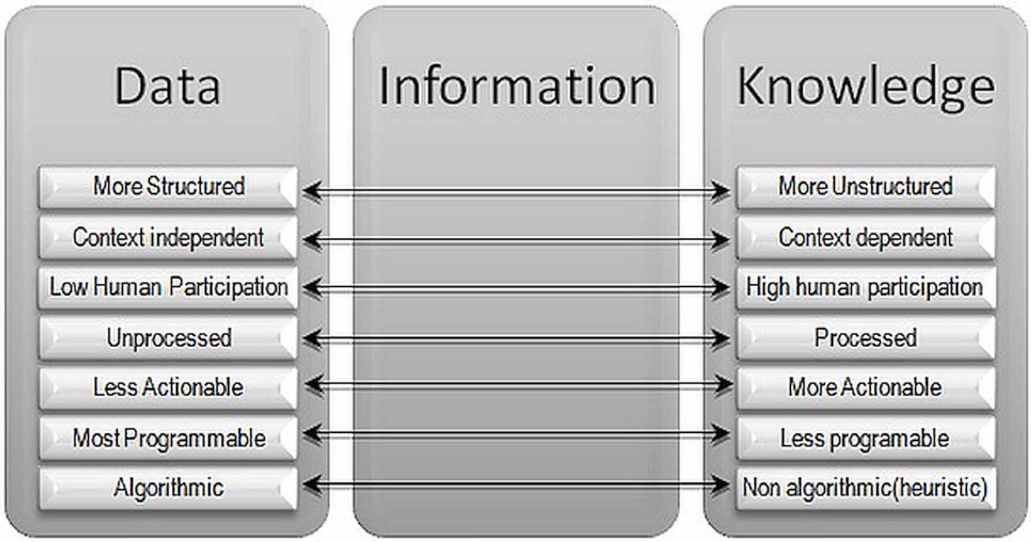
For understanding, a clear distinction between data and information could be done by using narrative that "Data are raw facts and figures that contain bits of information rather than information itself". These bits of information are not useful until they are processed, organized, interpreted, structured and converted into meaningful and useful form. The processed form of data which is converted into meaningful and useful form is known as "Information". The application of data and information is known as Knowledge that extends to appropriate collection of useful information. (Definition of Terms, 2002).

The term "knowledge" is part of our everyday language. It is used interchangeably in between knowhow and wisdom (Altaher, 2011; Davenport, Long & Beer, 1998). At one time, it is used as simple as knowhow about anything and at other time is used at its highest level as wisdom. Many times, it is identically used as "information" also, but actually there is a great amount of difference between them (Gamble & Blackwell, 2001). The perspective on data, information and knowledge is elaborated in Figure 1.

During the 1990s, the usage and application of information technology became increasingly common and the basic simple-minded view of the term "knowledge" was widely spread among the masses (Davenport, Long & Beer, 1998; Heier, Borgman, Hans & Manuth, 2005).

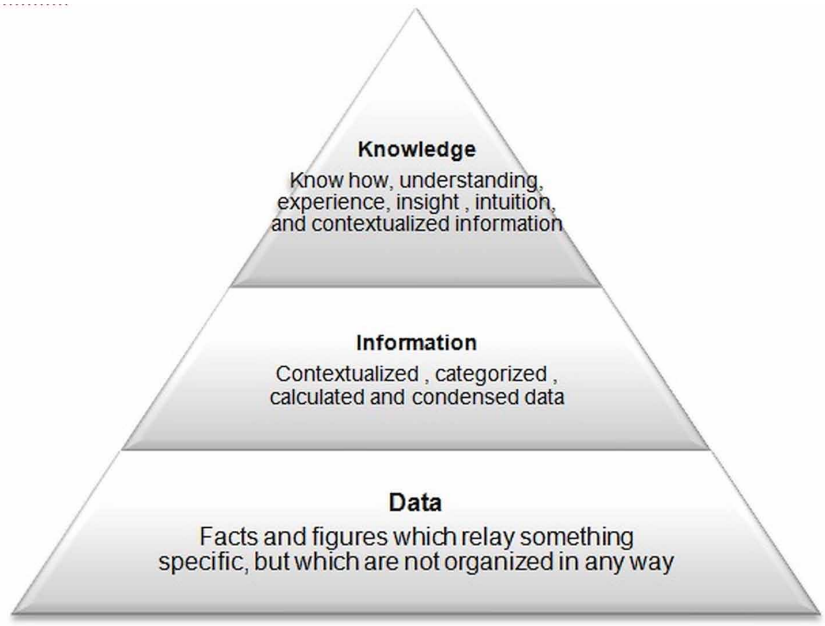
Figure 2 explains the difference between the three elements clearly.

Figure 1. Adapted from data, information and knowledge attributes (Davenport et al., 1998; Probst et al., 2000; Awad & Ghaziri, 2004)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Adapted from data, information and knowledge differentiation (Thierauff, 1999), (Davenport & Prusak 2000)



Different Types of Knowledge

Knowledge is considered as a mixture of grounded perception, contextual information, expert intuition, standards, and framed understanding that postulates a specific environment and structure for assessing and integrating new innovations in experiences and information (Nonaka, 1994; Frost, 2010). Today, knowledge is an integral part of organizations that is embedded in routines, practices, norms and repositories of organizations. In order for KM to succeed, one needs a deep understanding of what constitutes knowledge (Gamble & Blackwell, 2001; Frost, 2010). An important and crucial step in KM is to understand different forms of knowledge that exist and ability to differentiate between different types of knowledge (Razzaque & Jalal-Karim, 2010; Pasha, 2012).

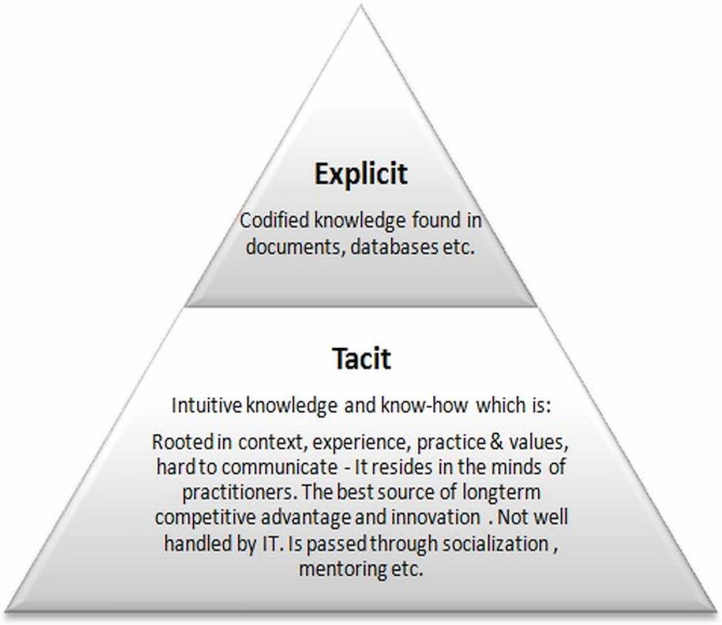
In KM with reference to business terms, knowledge is usually referred to as "Explicit knowledge" that exists in documents in form of codified knowledge and "Tacit knowledge", often mentioned as experience of individuals and is in non-codified form (Bukowitz & Williams, 1999). This division is explained in Figure 3.

The theory of organizational learning in relation to KM is originated and lies in correlation between Explicit and Tacit knowledge. Nonaka developed and explained this concept in the 1990s as a key foundation of this field. Explicit and Tacit knowledge types should not be considered as absolute points. However, for understanding knowledge, it is imperative to define both these terms distinctly (Botha, Kourie & Snyman, 2008).

The source of tacit knowledge is internal discrete processes of organizations that are stored in human beings as a result of their individual experiences and perceiving ability. Such type of knowledge obtained may be occasionally referred to as observation, experience, individual ability or internalization (Botha, Kourie & Snyman, 2008).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Adapted from (Brown &Duguid 1991, Cook & Brown 1999, Bukowitz & Williams 1999)



Explicit knowledge is normally stored on mechanical or highly technological devices such as electronic documents or databases. This form of knowledge is comparatively less useful until or unless shared among community that uses collaborative technology working together at any given time or place (Thierauf, 1999).

Some researchers make a further distinction and talk of embedded knowledge. In this way, one differentiates between knowledge embodied in people and that embedded in processes, organizational culture, routines, etc. (Woods, Cordata & Horvarth, 2000-01). Gamble & Blackwell (2001) used a scale consisting of represented-embodied-embedded knowledge, where the first two closely match the explicit-tacit. It has also been generally accepted that both the explicit and tacit components of organizational knowledge play an important role in innovation (Hall &Andriani, 2003).

The relationship between data, information, explicit knowledge, tacit knowledge and wisdom is shown in Figure 4.

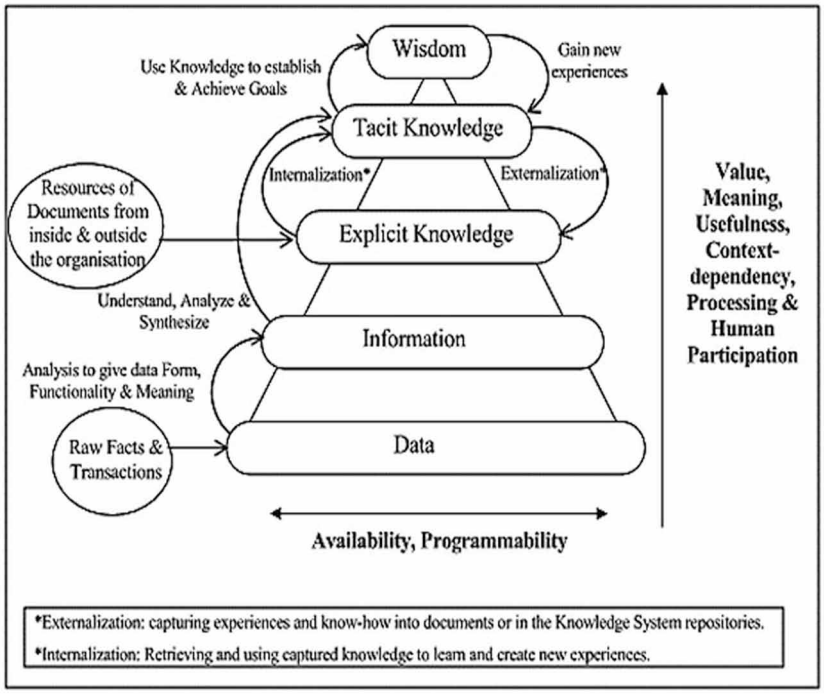
What is Knowledge Management?

Knowledge management is essentially about getting the right knowledge to the right person at the right time (Grace & Buttler, 2005). Knowledge management may also include new knowledge creation, or it may solely focus on knowledge sharing, storage, and refinement (Davenport, Long & Beer, 1998; Frost, 2010).

It is important to remember that knowledge management is not about managing knowledge for knowledge's sake. The overall objective is to create value and to leverage, improve, and refine the firm's competences and knowledge assets to meet organizational goals and targets (Botha, Kourie & Snyman, 2008; Jones, 2006).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Adapted from data, information, knowledge and wisdom (Davenport et al., 1998; Probst et al., 2000; Awad & Ghaziri, 2004; Bierly et al., 2000; NDR, 2003)



KNOWLEDGE MANAGEMENT IN HEALTHCARE

Overview

In the healthcare industry, the term "Health Informatics" or "Medical Informatics" is referred to the interdisciplinary study of resources and methods for the management of health information in healthcare services planning, management and delivery (NICHSR). (. This information is generated as a result of the production of large amounts of data (also known as "Big Data") in healthcare related studies that included electronic health records, medical health records, medical research, medical terminologies, methodologies, health information exchange, etc. (Rodger, 2015). Big Data in healthcare is not simply about the volume, velocity and variety of the data in storage, it is also about the potential value of the data, how it is processed and converted into information and how knowledge is perceived out of it (Rodger, 2015).

The healthcare industry sits on the hinge of a future in which doctors can instantly share imaging and test results with colleagues in the same building or across the city, country or continent (Persson et al., 2008). Patients should have access to their own medical history and be able to transfer it from one healthcare Centre to another. Innovation is a core competency of all healthcare institutions (Ruaranen, Lansisalmi, Kivimaki & Aalto, 2003). Further, implementation of new and enhanced IT systems and advanced genetic reengineering are drastically changing health care organizations, creating new opportunities for innovation and for the enhancement of the productivity of the actual processes (Leal-Rodríguez et al., 2013).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Innovation in healthcare continues to be a driving force in the quest to balance cost cuts and health care quality. Consequently, innovation has become a key piece for the achievement of a sustainable competitive advantage (O'Hara et al., 2013). In many instances, knowledge in the healthcare industry is mostly a blend of tacit and explicit knowledge which not only depends upon structured or direct resources of information but also is highly based on individuals' experiences and perceptions, which increase the difficulty of capturing and reusing it for innovation (Hussain & Raza, 2004). On the other hand, one of the other big issues could be "The culture of knowledge sharing and using it instinctively" (Abdul Rahman, 2011).

KM in Healthcare Industry of Pakistan

Pakistan is basically a developing country and is densely populated with more than 60% of the population living in rural areas. The rate of illiteracy is also very high and as time passes by, the population growth rate is also increasing day by day resulting in the increase of problems for healthcare industry. This requires a strong and effective healthcare industry to take care of healthcare problems of the people. Even though there are many well trained and qualified professionals and some state of the art healthcare centers present and scattered throughout the county, they still lack effective and quality knowledge management procedures (Mahmood, Burney, Abbas & Rizwan, 2012).

Importance of KM in Healthcare Industry of Pakistan

Pakistan is among those developing countries where the importance of knowledge management in healthcare services has been realized but its practices have failed to get sufficient attention. Lack of awareness about KM technologies and associated processes are among some common problems reported in recent studies. Also, the complexity involved in KM implementation has led to discouraging results reported in the research. One report shows about 84% of knowledge management projects fail. Thus, many Pakistani managers are reluctant to implement technology based KM in their organizations (Pasha, 2012).

Adoption of KM in Healthcare Industry of Pakistan

While knowledge management is becoming an established discipline with many applications and techniques, its adoption in healthcare has been challenging. Although the growth trend of this industry along with implementation of Knowledge Management Systems (KMS) have shown some notable achievements it still fails to meet the expectations of many with respect to efficiency and effectiveness (Pasha, 2012).

The lack of basic knowledge management procedures and processes also prevents the implementation of E-health, Tele-health or Tele-medicine in the country. These types of systems are generally required in those areas where the population lives far away from healthcare centers and unable to approach a proper healthcare service.

KM, ICT and Healthcare Industry

The use of Information and Communication Technologies (ICT) in health care is generally referred as e-Health, Tele-health or Tele-medicine:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Tele-Medicine:** Simply put, Tele-medicine is the provisioning of healthcare services at a distance using ICT. E-Health not only refers to the role of ICT in healthcare, but also to a way of thinking, a state of mind, and a commitment to improve the standard of health care being provided worldwide by using the ICT industry;
- **E-Health:** E-Health is the intersection of Information Sciences, Computer Science, Information Technology and Healthcare. It deals with various approaches used in storage, retrieval, and transfer of information in healthcare and biomedicine. It includes computers as well as clinical guidelines, formal medical terminologies, and information and communication systems (Aqil, Zain, Nadeem & Qamar, 2013). The successful implementation of all these systems requires an effective and quality Healthcare Knowledge Management System.

KNOWLEDGE MANAGEMENT PROCESS AND ITS IMPORTANCE

Overview

A solid, comprehensive and integrated KM process framework is the most important resource and asset for an organization to implement Knowledge Management Systems effectively and efficiently and for getting maximum benefit of their KM Strategies. An assimilated KM Process Framework is a step-by-step guideline that provides effective standard operating procedures to manage knowledge assets of an organization for gaining maximum cost and time effective benefit (Arshad, Noordin & Othman, 2016).

Although much research has been accomplished regarding KM frameworks in various fields of studies, the healthcare industry of Pakistan is still lacking a considerable amount of research and literature with respect to effective knowledge management. Still some of the researchers have been striving to conduct true research in this domain. In our opinion, one of the basic reasons for such circumstances is an intersection of the 3 domains of Healthcare, Knowledge Management and Information Technology (Arshad, Noordin & Othman, 2016).

Existing Work

In Table 1, the authors have summarized some of the work done in this domain by researchers related to KM in the healthcare industry of Pakistan. Each has its own Pros and Cons but none of them has presented a KM framework that effectively deals with procedural implications of knowledge management in the healthcare industry of Pakistan (Arshad, Noordin & Othman, 2016).

CASE STUDY – DESCRIPTION

Knowledge Management Practices at Public Sector Healthcare Organizations (Hospitals) in Pakistan

This section provides an excerpt of information from a case study conducted on 2 government hospitals (A & B–pseudonym used to ensure confidentiality) in one of the major provincial capital city in Pakistan. These are among the largest and oldest public healthcare organizations governed and controlled

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Adopted - A review of Knowledge Management Frameworks proposed by some researchers for Healthcare Industry of Pakistan (Arshad, Noordin, Othman, ICT4M2016)

Framework	Description
Fehmida & Ali et al., 2004	**Knowledge Management Framework To Operationalize Experiential Knowledge** • Mapping of explicit form of knowledge with tacit form of knowledge • Objective of this model is that explicit knowledge in the form of Clinical practice guidelines (CPG) is enriched by tacit, experiential knowledge found in the form of documented experience.
Pasha et al., 2012	**A Theoretical Technology Driven Knowledge Management Framework** • Aims to advance the understanding of the key concepts of KM and the relationship between available technologies and Knowledge Processes • Identifies key factors for making a KM initiative successful
Mahmood, Nadeem, et al, 2012	**Healthcare Workflow Management and ICT Model** • Process of management of knowledge involves the ability to identify and capture existing knowledge, create new knowledge and transfer existing knowledge throughout the organization. • Mainly emphasizes on the flow of knowledge from one entity to another.
Jalal, Karim et al, 2010	**The HC KM Model** • Knowledge representation is referred to as storing and processing of information and knowledge so that it can be used by other applications.
Hameed, Shafqat et al, 2012	**The Enhanced HC KM Model** • Enhanced this model by introducing the personal identity (national identity) of a patient to be recorded by the system.

by the local government and comprising of more than 2000 beds collectively. Hospital "A" employs 1,025 healthcare professionals with a total count of 1800 employees and staff in its 36 departments and is one of the oldest public-sector healthcare organizations in Southeast Asia. Hospital B employs 1,225 healthcare professionals with a total count of 2,100 employees and staff in its 32 departments with its roots being older than 70 years. The hospitals seem to be overcrowded with a significant number of in-patients (A > 30000, B > 25000 per year) and out-patients (A > 1000000, B > 900000 per year).

The case study was conducted to obtain an account of information on practical grounds related to existence and implementation of KM processes in the top and famous public sector Pakistani healthcare organizations.

The researchers selected and used a purposive sampling technique to answer the research questions (see Table 2). The researchers collected data and information for their study by help of generalized interviews, structured walkthroughs, existing document reviews and self-observations. While performing their exploratory study, the authors also collected information regarding the KM dimensions (organizational culture, leadership, technology and politics) that have a critical impact on the success of KM implementation in healthcare industry. A substantial number of people including healthcare professionals, IT staff, support staff and some educated patients were met to dig out the information (see Table 3).

The interviews and data collection objectives of the authors were generally based upon the major domains and constructs shown in Table 4.

Findings

The overall findings and results helped the authors to propose an integrated and interoperable KM process framework for healthcare. Table 5 provides the level of agreement of respondents based upon the questionnaire constructs that were conducted by help of interviews.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Case demographics

	Organization Case A	Organization Case B
Status	Public Sector	Public Sector
Type	Healthcare Organization	Healthcare Organization
Beds	1055	945
Employees	1800	2100
Departments	36	32
Inpatients (Per Year)	>30000	>25000
Outpatients (Per Year)	>1000000	>900000

Table 3. Population sampling

Population Sampling	Organization Case A				Organization Case B			
	Participants				Participants			
	Total	Male	Female	Age Group	Total	Male	Female	Age Group
Management	9	6	3	40 – 65	11	7	4	40 – 65
Healthcare Professionals	23	14	9	29 – 45	21	15	6	29 – 45
IT Staff	10	7	3	23 – 43	9	7	2	23 – 43
Support Staff	13	5	8	22 – 37	10	3	7	22 – 37
Educated Patients	15	12	3	33 - 54	11	10	1	33 - 54
Total	70	44	26		62	42	20	
% age	100%	62.85%	37.15%		100%	67.74%	32.25%	

Table 4. Questionnaire constructs

1	Know how about the terminology “Knowledge Management”
2	Know how about the importance of Knowledge Management
3	Role of KM in healthcare
4	Benefits of Knowledge Management in healthcare
5	Organizational Culture
6	Financial and top management support for KM implementation
7	Implementation of KM processes in the organization
8	Knowledge innovation, sharing, usage and storage
9	Existence of Knowledge Management Strategy and framework
10	Role of KM in decision making
11	Best practices about KM for patient care
12	Organizational and Technology infrastructure

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Level of agreement and responses by the respondents to questionnaire constructs (summarized)

Case Organizations A & B							
Summary of Question Constructs and Responses n = (Out of 132 Respondents)	**Level of Agreement**	**1**	**2**	**3**	**4**	**5**	**6**
		Don't Know n(%)	**Strongly Disagree n(%)**	**Somewhat Disagree n(%)**	**Neutral n(%)**	**Somewhat Agree n(%)**	**Strongly Agree n(%)**
Have knowhow about the term "knowledge management" (KM)		15.71%	7.14%	5.71	10%	54.28%	7.14%
Have knowhow about the importance of KM		15.71%	7.14%	5.71%	10%	54.28%	7.14%
Have awareness about role of KM in healthcare		28.57%	5.71%	5.71%	7.14%	37.71%	17.14%
Aware of benefits of KM in healthcare		28.57%	0%	0%	14.28%	38.57%	18.57%
KM process framework supported by organizational culture		52.85%	12.85%	11.42%	21.42%	1.42%	0%
KM processes implemented in the organization		28.57%	18.57%	24.28%	27.14%	1.42%	0%
KM process framework implementation supported by top management		47.14%	4.28%	5.71%	8.57%	21.42%	12.85%
Financial support available for KM process implementation		28.57%	2.85%	4.28%	7.14%	24.28%	32.85%
Knowledge innovation, usage, sharing and storage supported by organizational culture		28.57%	5.71%	0%	15.71%	28.57%	21.42%
Existence of KM strategy and framework		57.14%	7.14%	4.28%	21.42%	5.71%	4.28%
Role of KM in decision making		52.85%	0%	0%	4.28%	11.42%	31.42%
Implementation of best practices about KM for patient care		48.57%	5.71%	8.57%	31.42%	4.28%	1.42%
KM supported by organizational infrastructure		45.71%	0%	4.28%	2.85%	8.57%	38.57%
Existence of sufficient Technology Infrastructure		37.71%	2.85%	7.14%	2.85%	15.71%	37.71%

The following points were highlighted as findings after the collected data was analyzed and processed. As per the findings of study:

1. Both the organizations were lacking the basic KM process framework implementation;
2. The results of the discussions and observations showed, that most of the professionals and related stake holders from both the healthcare organizations do understand the term Knowledge Management, its importance and benefits for healthcare but none of them succeed to satisfy the researchers about their knowledge of existence of any KM process framework that has been professionally and competently implemented;
3. Despite the fact that both the organizations do possess a sufficient and competent IT infrastructure along with sufficient funds and top management support, none of the knowledge workers were aware of implementation of any professional KM process framework to support decision making processes and operations;
4. As both the cases under study were public sector organizations, it was found that both were working under the strong influence of political framework that was not possible to be further investigated;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. The researchers also found that both the organizations lacked the basic culture of information sharing and usage among knowledge workers and the organizational culture was critically impacted by a strong influence of social culture and other internal and external factors;
6. The researchers also discovered that most of the new knowledge creation and innovation takes place in discussions held between the healthcare professionals over the table talks and tea breaks during the working hours;
7. The healthcare professionals were not aware of any sort of existence of a system or framework which can help them or the other stakeholders to use and disseminate knowledge;
8. There was no concept of Tele-medicine or E-health being used in support of patients;
9. Furthermore, by the help of structured walkthroughs, document reviews and existing system review, it was established that both the healthcare organizations do lack a proper and competent Knowledge Management process framework to support daily based knowledge management activities.

Based upon above case study and findings, the researchers strive to define and propose a competent and completely integrated KM Process Framework to implement in Pakistani Healthcare industry and to support daily based KM operations and activities.

THE PROPOSED KM PROCESS FRAMEWORK

Based upon facts and findings explored during the study that was supported by interviews, structured walkthroughs, documentary reviews and self-observatory process, the researchers strive to propose an integrated and comprehensive process framework of KM for healthcare organizations in Pakistan and represent the main components of the KM process model to facilitate its understanding and to show the relationships among the different phases. The proposed KM process framework consists of ten comprehensive phases as shown in Figure 5, of which each is having its own importance and implications:

Phase 1 starts the KM process by identifying and defining the specific needs and requirements of different types of knowledge for the healthcare organizations e.g. content requirements related to type and format of data and information to be stored.

Phase 2 refers to the identification and segregation of specific type of knowledge to fulfill the needs of healthcare organizations e.g. classification of knowledge into different types.

Phase 3 refers to the identification and definition of all the resources required for knowledge management e.g. technical resources, human resources, financial resources etc.

Phase 4 refers to the identification and definition of influential factors that may affect activities and components of the KM process, and deciding the required procedures and methods to successfully deal with these factors e.g. organizational culture, political culture, social culture, etc.

Phase 5 refers to all the areas that are related to Acquisition, creation or identification of knowledge related processes, procedures, policies & resources e.g. rules and regulations, legal obligations, etc.

Phase 6 refers to deciding, sorting and sequencing of processing procedures and activities to handle and manage all types of knowledge and knowledge management resources.

Phase 7 helps choosing and applying the required ICT system specifications and IT tools that enhance and support the KM activities e.g. hardware requirements, software requirements, etc.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Phase 8 describes the procedures related to retrieval, application, generation, innovation & sharing of new knowledge e.g. the methods and approaches to be used for knowledge production and communications, etc.

Phase 9 is related to the usage of knowledge in an effective and efficient way to gain maximum advantage, e.g., methods for implementation and application of knowledge perceived.

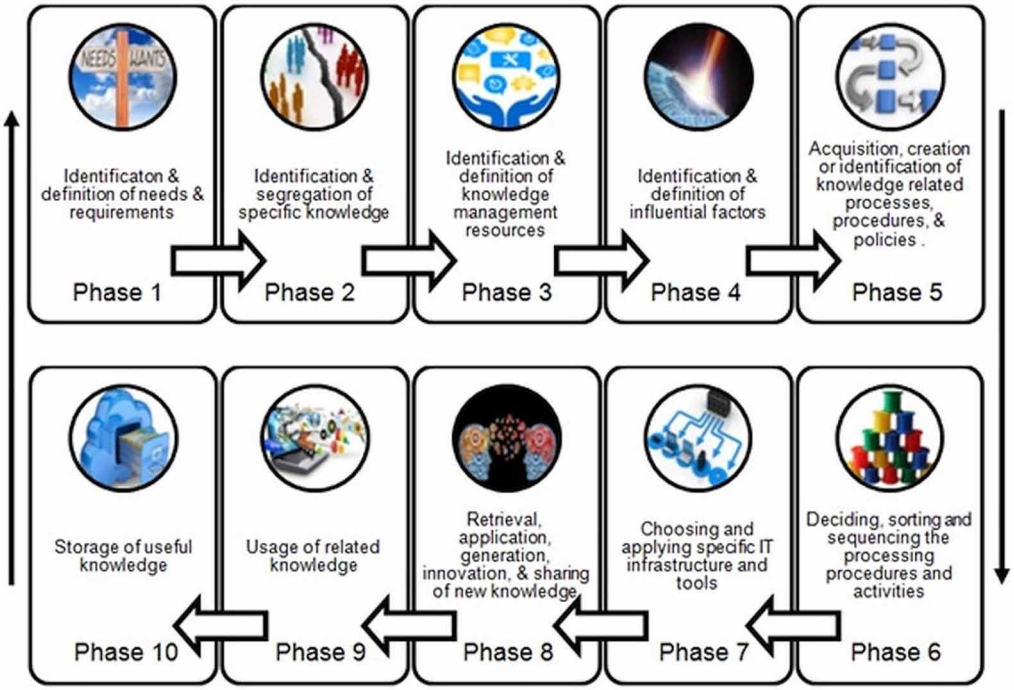
Phase 10 seeks to identify suitable methods and procedures for storage of new knowledge, e.g., effective processes to store useful knowledge in databases in specific format.

Each of the phases defined above has its own importance and implications for knowledge management in the healthcare industry of Pakistan. Collectively, all the phases can be summarized into a compact form which could be divided into 4 broad categories. Figure 6 explains the proposed KM process framework in a summarized way.

The 4 broad categories in which the 10 phases of Knowledge Management Process Framework could be summarized are:

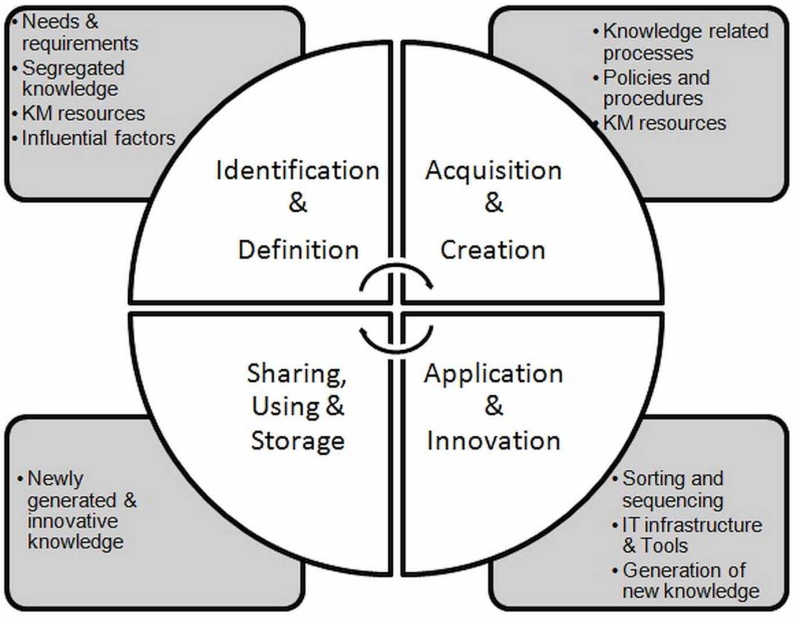
1. **Identification & Definition:** Comprises of needs, requirements, segregated knowledge, resources and influential factors;
2. **Acquisition & Creation:** Comprises of knowledge related processes, policies and procedures and KM resources;
3. **Application & Innovation:** Comprises of sorting & sequencing, IT infrastructure & tools, and generation of new knowledge from existing knowledge;
4. **Sharing, Using & Storage:** Comprises of newly generated and innovative knowledge.

Figure 5. The proposed KM process framework adapted from – (Arshad, Noordin & Othman, 2016)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Adapted - summarized representation of proposed knowledge management process framework (Arshad, Noordin, Othman, 2016)



The four broad categories are hereby summarized and represented in Figure 6.

KIIPF – KM INTEGRATED INTEROPERABLE PROCESS FRAMEWORK

The authors here emphasize the importance that healthcare is directly related with lives of people. A healthcare professional must have proper and accurate knowledge before dealing with his patients and providing advice or treatment. The importance of Knowledge Management in healthcare can be understood in terms of "right knowledge for the right treatment is a matter of life or death." One must be very specific and selective in this respect.

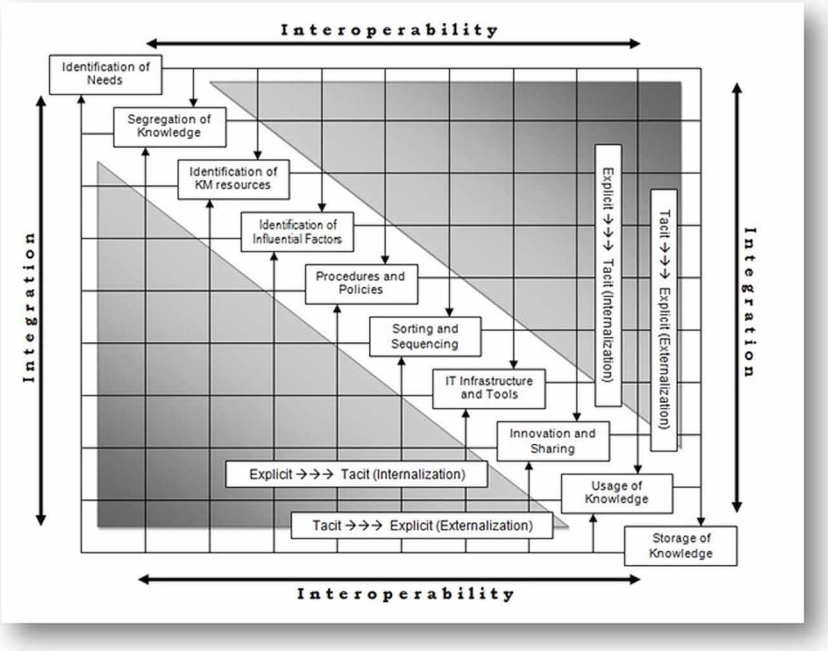
Acquiring the right knowledge depends on how the existing knowledge is managed and how the new knowledge is created with the help of innovative ideas and processes. A proper and scientific knowledge management process framework can make sure the availability of right knowledge to the right person at the right time. This may save many lives.

The KM Process Framework defined in Figure 5 can further be extended and elaborated in terms of Knowledge Management Integrated & Interoperability Process Framework (KIIPF) Grid as shown in Figure 7.

The KIIPF Grid explains clearly how the knowledge management process phases can work together to enhance integration and interoperability among them. This will increase the effectiveness and efficiency of the process framework. Each phase is inter-connected with all other phases in the grid to create interdependency, hence dealing with the sensitivities and margins of errors among each phase. Deficiency in one phase is ultimately going to affect the working of all other phases.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. The KIIPF – knowledge management integrated interoperability process framework grid



The last 3 phases in the suggested process framework define the conversion of existing knowledge into new and innovative knowledge, its sharing and usage and in the end storage of the newly created useful knowledge. Hence, their integration and interoperability works mainly for internalization (Explicit to Tacit conversion of knowledge) and externalization (Tacit to Explicit conversion of knowledge) for storage (Nonaka, 1994).

CONCLUSION

A comprehensive analysis emphasizes that there is always an opportunity to improve and enhance the way the knowledge is handled in healthcare organizations. Due to technological innovations and procedural improvements, healthcare organizations in Pakistan are striving to expand and increase dependency on integrated Knowledge Management Process Models and Frameworks that not only improve the effectiveness and efficiency of Healthcare Information Systems in Healthcare Organizations of Pakistan, but also save time and effort to manage knowledge.

The healthcare sector can address the complexity and challenges of KM by adopting practical and effective KM strategies which help to improve the quality of care for patients by providing them continuity and consistency of care. The current state of KM in healthcare is progressing and improving which is a positive sign for healthcare professionals and administrators to deal with pertinent challenges of KM implementation by collaborating and sharing useful knowledge and relying on the best knowledge available. This also helps open new horizons for innovations related to KM and its implementation in Healthcare sector.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In this study, the authors suggested an integrated interoperable comprehensive phased Knowledge Management Process Framework which helps and guides to pave ways for all the stakeholders in the Healthcare industry of Pakistan to share and use the combined knowledge in most beneficial way increasing competitive advantage and saving valuable resources. The Knowledge in a Healthcare Organization is an asset which must not be compromised upon, so must be managed carefully. The authors believe that the current state of healthcare can be improved by implementing such KM systems that help to increase quality of healthcare by increasing efficiency and effectiveness of KM implementation in healthcare. The outcome results into practicing better evidence based medicine by innovation, collaboration and sharing of knowledge in a proficient and effective way.

It has been mentioned earlier that the case studies were performed on public sector healthcare organizations which work under the control and influence of local government bodies. A direct interference and involvement to implement, evaluate, verify and validate the usefulness of the proposed KM process framework requires government support and authorization without which the study is limited to proposition.

For future and further research, the researchers plan to develop an "Integrated KM Model for Healthcare Industry in Pakistan" based on this KM process framework and evaluate the applicability of the proposed model. After that, a survey will be conducted to evaluate the effectiveness of the KM Model in the healthcare industry of Pakistan.

REFERENCES

Abdul Rahman, R. (2011). Knowledge sharing practices: A case study at Malaysia's healthcare research institutes. *Journal of International Information and Library Review*, *43*(4), 207–214. doi:10.1080/10572317.2011.10762902

Akhavan, P., Jafari, M., & Fathian, M. (2005). Exploring failure-factors of implementing knowledge management systems in organizations. *Journal of Knowledge Management Practice, 6*, 1-8.

Akram, M., & Khan, F. J. (2007). Healthcare services and government spending in Pakistan (Working Paper). *Pakistan Institute of Development Economics (PIDE)*.

Alavi, M., & Leidner, E. D. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *Management Information Systems Quarterly*, *25*(1), 107–136. doi:10.2307/3250961

Allameh, S. M., Zare, S. M., & Davoodi, S. M. R. (2011). Examining the impact of KM enablers on knowledge management processes. *Procedia Computer Science*, *3*, 1211–1223. doi:10.1016/j.procs.2010.12.196

Altaher, A. M. (2011). Knowledge management process implementation. *International Journal of Digital Society*, *1*(4), 265–271. doi:10.20533/ijds.2040.2570.2010.0032

Aqil, B., Zain, A., Nadeem, M., & Qamar, A. (2013). Prospects for mobile health in Pakistan and other developing countries. *Scientific Research, Advances in Internet of Things, 3*(2A).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Arshad, A., Noordin, M. F., & Othman, R. (2016). A comprehensive knowledge management process framework for healthcare information systems in healthcare industry of Pakistan. In *Proceedings of the 6th International Conference on Information and Communication Technology for the Muslim World (ICT4M), IEEE Conference Publications* (pp. 30-35). 10.1109/ICT4M.2016.019

Bhutto, R. A., & Maqsood, A. (2007). Evaluating knowledge & information-based healthcare system: A case study of Dow University of health sciences & civil hospital Karachi.

Botha, A., Kourie, D., & Snyman, R. (2008). *Coping with Continuous Change in the Business Environment, Knowledge Management and Knowledge Management Technology*. London: Chandice Publishing Ltd. doi:10.1533/9781780632056

Bukowitz, W., & Williams, R. (1999). *The knowledge management fieldbook*. London, UK: Financial Times/Prentice Hall.

Davenport, T., De Long, D., & Beer, M. C. (1998). Successful knowledge management projects. *Sloan Management Review*, *39*(2), 43–57.

Davenport, T. H., & Prusak, L. (2000). *Working knowledge: How organizations manage what they know*. Boston, MA: Harvard Business School Press.

Definition of Terms. (2002). Knowledge Based Solutions. Retrieved June 13, 2016, from http://www.kwbsolutions.com/kbsterms.htm

Fauzan Noordin, M. (2009). *ICT and Islam*. Kualalumpur, Malaysia: IIUM Press International Islamic University Malaysia.

Fauzan Noordin, M., Othman, R., & Azura Zakaria, N. (2011). Peopleware and heartware – The philosophy of knowledge management, research and innovation in information systems. In *Proceedings of the International Conference on Research and Innovation in Information Systems*.

Frost, A. (2010). *KMT – An educational KM site*. From http://www.knowledge-management-tools.net

Gamble, P. R., & Blackwell, J. (2001). *Knowledge management: A state of the art guide*. London: Kogan Page Publishers Ltd.

Govindarajan, V. (2007). The importance of strategic innovation, from http:/www.vijaygovindarajan.com/2006/03/the_importance_of_strategic_in.htm

Grace, A., & Butler, T. (2005). Beyond knowledge management: Introducing learning management systems. *Journal of Cases on Information Technology*, *7*(1), 53–70. doi:10.4018/jcit.2005010104

Hall, R., & Andriani, P. (2003). Managing knowledge associated with innovation. *Journal of Business Research*, *56*(2), 145–152. doi:10.1016/S0148-2963(01)00287-9

Hameed, S., Jawad, K., & Kashif, M. (2012). Effectual dynamics and prolific usage of knowledge management & engineering in healthcare industry. *Life Science Journal, 9*(2).

Heier, H., Borgman, H. P., & Manuth, A. (2005). Siemens: Expanding the knowledge management system ShareNet to research & development. *Journal of Cases on Information Technology*, *7*(1), 92–110. doi:10.4018/jcit.2005010106

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hussain, F., & Ali Raza, S. (2004). A knowledge management framework to operationalize experiential knowledge: mapping tacit medical knowledge with explicit practice guidelines. In *Proceedings of the National Conference on Emerging Technologies*.

Jones, K. (2006). Knowledge management as a foundation for decision support systems. *Journal of Computer Information Systems*, *46*(4), 116–124.

Leal-Rodríguez, A., Leal-Millán, A., & Roldán-Salgueiro, J. L., & Ortega-Gutiérrez, J. (2013). Knowledge management and the effectiveness of innovation outcomes: The role of cultural barriers. *Electronic Journal of Knowledge Management*, *11*(1), 62–71.

Mahmood, N., Burney, A., Abbas, Z., & Rizwan, K. (2012). Data and knowledge management in designing healthcare information systems. *International Journal of Computers and Applications*, *15*(2), 34–39. doi:10.5120/7745-0798

NICHSR – National Information Center on Health Services Research and Healthcare Technology. (n. d.). Informatics. Retrieved from https://www.nlm.nih.gov/hsrinfo/informatics.html

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, *5*(1), 14–37. doi:10.1287/orsc.5.1.14

Nonaka, I., Takeuchi, H., & Umemoto, K. (1996). A theory of organizational knowledge creation. *International Journal of Technology Management*, *11*(7-8), 833–845.

OHara, K., Esteves, D., Brás, R., Rodrigues, M., Rodrigues, R., & Pinheiro, P. (2013). Health Learning Practices in Adolescents Using Physical Activity. *Journal of Cases on Information Technology*, *15*(1), 83–97. doi:10.4018/jcit.2013010106

Pasha, M. A., & Pasha, S. (2012). A pragmatic approach for implementing knowledge management in Pakistani organizations using open source. *International Journal of Computers and Applications*, *49*(7), 10–18. doi:10.5120/7638-0721

Persson, A., Stirna, J., & Aggestam, L. (2008). How to disseminate professional knowledge in healthcare: The case of Skaraborg hospital. *Journal of Cases on Information Technology*, *10*(4), 41–64. doi:10.4018/jcit.2008100103

Ratnapalan, S., & Uleryk, E. (2014). Organizational Learning in Health Care Organizations. *Systems*, *2*(1), 24–33. doi:10.3390ystems2010024

Razzaque, A., & Jalal-Karim, A. (2010). Conceptual healthcare knowledge management model for adaptability and interoperability of HER. In *Proceedings of the European, Mediterranean & Middle Eastern conference on Information Systems*.

Razzaque, A., & Jalal-Karim, A. (2010). The influence of knowledge management on EHR to improve the quality of healthcare services. In *Proceedings of the European, Mediterranean & Middle Eastern Conference on Information Systems*.

Rodger, A. J. (2015). Discovery of medical big data analytics: Improving the prediction of traumatic brain injury survival rates by data-mining Patient Informatics Processing Software Hybrid Hadoop Hive. *Informatics in Medicine Unlocked*, *1*, 17–26. doi:10.1016/j.imu.2016.01.002

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ruaranen, R., Lansisalmi, H., Kivimaki, M., & Aalto, P. (2003). Innovation in healthcare: A systematic review of recent research. *Nursing Science Quarterly*, *19*(1), 66–72. PMID:16407603

Thierauf, R. J. (1999). *Knowledge management systems for business*. USA: Greenwood Publishing Group.

Woods, A. J., Cordata, J., & Horvarth, J. A. (2000-2001). *The Knowledge Management Yearbook 2000-2001*.

World Health Organization (WHO). (n. d.). Everybody's business: Strengthening health systems to improve health outcomes, WHO's Framework for Action. Retrieved from http://www.who.int/health-systems/strategy/everybodys_business.pdf

This research was previously published in the Journal of Cases on Information Technology (JCIT), 19(3); edited by Andrew Borchers; pages 24-41; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 9
Semantic Health Mediation and Access Control Manager for Interoperability Among Healthcare Systems

Abdullah Alamri
University of Jeddah, Saudi Arabia

ABSTRACT

Healthcare systems have evolved to become more patient-centric. Many efforts have been made to transform paper-based patient data to automated medical information by developing electronic healthcare records (EHRs). Several international EHRs standards have been enabling healthcare interoperability and communication among a wide variety of medical centres. It is a dual-model methodology which comprises a reference information model and an archetype model. The archetype is responsible for the definition of clinical concepts which has limitations in terms of supporting complex reasoning and knowledge discovery requirements. The objective of this article is to propose a semantic-mediation architecture to support semantic interoperability among healthcare organizations. It provides an intermediate semantic layer to exploit clinical information based on richer ontological representations to create a "model of meaning" for enabling semantic mediation. The proposed model also provides secure mechanisms to allow interoperable sharing of patient data between healthcare organizations.

1. INTRODUCTION

Over the last decades, healthcare systems have evolved to become more patient-centric. Healthcare is moving towards supporting a continuous medical process by including multiple healthcare professionals and institutions, utilizing ubiquitous computing healthcare environments and embracing technological advances (Martínez-Villasenõr et al., 2016). Nowadays, it is hard to believe that there is still a prevalence of paper-based patient data throughout most of the health sector, despite all efforts to automate and share

DOI: 10.4018/978-1-7998-1204-3.ch009

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

electronic health information among practitioners, patients and other stakeholders. Many efforts have been made to transform paper-based patient data to automated medical information by developing Electronic Healthcare Records (EHRs) and standards in order to enable the reuse and sharing of information amongst healthcare systems (Martínez-Villasenõr et al., 2016; Begoyan, 2007).

EHR systems provide a collection of patient health information which includes medications, laboratory test results, demographic data, progress notes, past medical history, radiology data and immunizations, among other health-related data. Some of the advantages of an EHR are digitalized information commonly implies, such as improving the accuracy of medical terminology and availability of relevant and necessary information at the point of care (Ambinder, 2005). Moreover, EHR provides a complete view of a patient's health record which, logically, supports other analytical health-based activities, which in turn will improve the decision-making process and electronic communication. It is important that healthcare systems cooperate and share information in order to improve the quality of healthcare by reducing errors in medical decisions, decreasing health costs, and enhancing patient care.

Interoperability between EHRs plays a significant role in healthcare development as it can deliver more efficient and effective patient care, and helps with the retrieval and processing of patients' clinical information from and to different health databases. A number of approaches to improve the interoperability of health systems have been applied in recent years. Several international EHRs standards have been enabling healthcare interoperability and communication among a wide variety of medical centres. It is a dual-model methodology which comprises of a reference information model and an archetype model. The archetype is responsible for the definition of clinical concepts which has limitations in terms of supporting complex reasoning and knowledge discovery requirements (Lezcano et al., 2011).

Healthcare information is very complex and therefore it is difficult to establish interoperability between EHRs. Hospital health records are extremely heterogeneous. Patient's health data tends to be recorded in different forms and formats. Patient's health information may include different types of recorded data: laboratory test results, treatment notes and medications, progress notes, referrals, imaging, medical charts, nursing notes etc. Healthcare information is stored in different proprietary formats, and in structured formats including databases, and unstructured documents. This lack of consistency leads to severe interoperability problems.

Several scientific researchers have suggested using semantic web technologies and ontology to overcome the aforementioned problems (Kolias et al., 2014; Cannoy and Iyer, 2008). The semantic web technologies provide enhanced capabilities that allow data to be processed in a more effective and accurate way, create the framework for interoperability between healthcare systems and integrate data from various sources by means of their semantic meaning. The interoperability of healthcare systems is defined in the Semantic-health report as "the ability, facilitated by applications and systems, to exchange, understand and act on citizens/patients and other health-related information and knowledge among linguistically and culturally disparate health professionals, patients and other actors and organizations within and across health system jurisdictions in a collaborative manner" (Lezcano et al., 2011; Stroetman et al., 2009).

The motivation to increase the use of EHRs is grounded in evidence that they may improve the quality, efficiency, safety, and patient satisfaction with care. However, the adoption of EHR alone is insufficient to realize the full promise of health improvement. To enable patients to keep track of their health data wherever they receive care, attention is now on health information exchange (HIE), defined as the reliable and interoperable electronic sharing of patient data between healthcare organizations, clinical information among physicians, nurses, pharmacists, other health care providers, and, health data repositories (Hersh et al., 2015).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The objective of this paper is to propose a semantic-mediation architecture to support semantic interoperability among healthcare organizations. This paper does not intend to propose a new EHR standard. Instead, it provides an intermediate semantic layer to exploit clinical information based on richer ontological representations in order to create a "model of meaning" for enabling semantic mediation; and to deal with heterogeneity across data sources in EHRs which has yet to be adequately addressed. The proposed model also provides secure mechanisms to allow interoperable sharing of patient data between the healthcare organizations. This paper addresses the following questions:

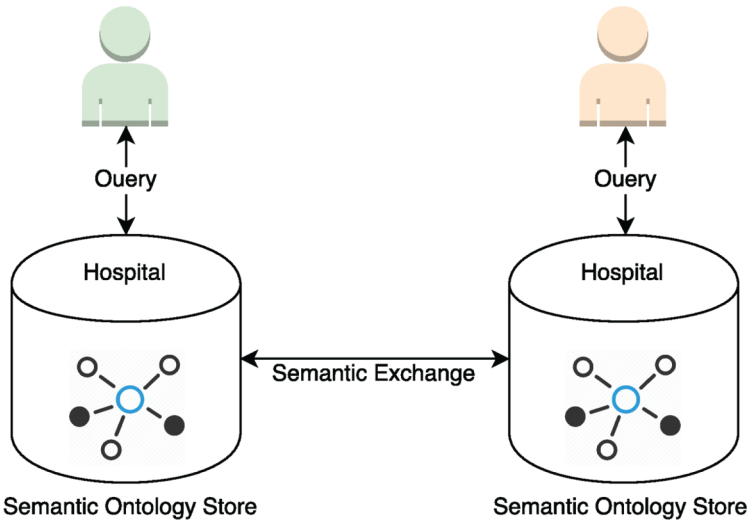
- How can healthcare information be represented using a semantic knowledge base?
- How to create a model that can deal with the issue of heterogeneity?
- How can requests for sensitive patient data disclosure be characterized and modelled?

The rest of the paper is organized as follows: Section 2 presents some background information about the essential aspects of health informatics related to electronic healthcare records (EHRs) and database design. It also presents some background information about the components of a semantic ontology knowledge base. In Section 3, we describe semantic health mediation and the access control manager. The following sections describe the main components of the model architecture. Section 4 presents a motivation example. The paper concludes with a number of suggestions for future research and development.

2. BACKGROUND

This section presents information about the essential aspects of health informatics related to electronic healthcare records (EHRs) and database design. Also, we offer some background information about the components of a semantic ontology knowledge base.

Figure 1. Interoperation Semantic Healthcare Systems



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.1. Electronic Health Records

EHR systems facilitate the management and sharing of patients' health information among health care providers. EHRs collect patient health records related to medications, diagnoses, hospital admissions, operations, imaging, laboratory tests, and pathology data. The main benefit of an EHR is that it improves the quality of health care delivery. Since the EHR concept seems to be complex, companies commercializing EHR have been implementing standards. A number of international organizations for standardization that have provided standard solutions for EHR are:

- International Organization for Standardization;
- European Committee for Standardization;
- Health Level Seven accredited by American National Standards Institute in the US.

International Organization for Standardization (ISO): The International Organization for Standardization (ISO) established ISO's Technical Committee ISO/TC 215. This committee introduced several standards related to health informatics. One of these standards, (ISO/TR 20514), defines the structure and context of a basic/generic HER (ISO, 2005). They also published ISO 13606-reference model which used to enhance EHR communication (Martinez-Costa et al., 2014; ISO, 2003).

European Committee for Standardization (CEN) in Europe: the European Committee for Standardization of Health Informatics (CEN/TC 251) presented a reference model, an archetype interchange specification, reference archetypes, security features, and exchange models. Also, CEN/ISO 13606 standard is a European standard approved by ISO designed to achieve semantic interoperability in electronic health record communication. The ISO/CEN 13606 standard is based on a dual-model architecture: archetypes to provide semantic meaning and the reference model structure (CEN, 2008; Lozano-Rubí et al., 2016).

Health Level Seven (HL7) in the US: the term 'HL7' refers to both the organization and to a set of messaging standards. HL7 are the health-messaging standards that support two messaging protocols: HL7 Version 2 and HL7 Version. HL7 Version 3 has developed the Clinical Document Architecture (CDA) for exchanging health data such as progress notes, discharge summaries, and results of physical examinations across healthcare systems (Benson and Grieve, 2016; Mandel et al., 2016; Sinha et al., 2012).

EHR systems have three architectures: basic, universal and distributed. In the basic HER architecture, a centralized EHRs database is developed in one organization (i.e. a hospital) and collects data from all healthcare systems operating in the hospital such as the laboratory system, radiology information system and others. In the universal HER architecture, a centralized EHR database is developed from a number of national/regional EHRs databases. In the distributed architecture, each healthcare organization (i.e. a hospital) has its own EHR with its own data model and with its own terminology standards (El-Sappagh et al., 2012).

2.2. Semantic Ontology Knowledge Base

The Semantic Web is a vision for the future of the World Wide Web where information is given well-defined meaning, providing logical connections of terms. In terms of the Semantic Web, the W3C recommends an RDF (Resource Description Framework) for semantic interoperability (Berners-Lee et al., 2001; Beckett and McBride, 2004; Turhan, 2011; Allemang and Hendler, 2011). RDF provides a powerful triple-based representation language for Universal Resource Identifiers (URIs). RDF Schema

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(RDFS) provides a data-modelling and structured vocabularies for RDF data. Web Ontology Language (OWL) provides greater expressivity in the objects and relations for describing domain knowledge than RDFS (McGuinness et al., 2004; McBride, 2004). The Semantic Web adopted description logic (DL) as a formal knowledge representation for the ontology language. The DL offers a good trade-off between expressivity, complexity and efficiency in knowledge representation, and reasoning about structured knowledge.

A DL ontology has two conceptual components:

- **TBox:** Terminological part;
- **ABox:** Assertional part.

The TBox is used to describe a conceptualization, assertions about classes, assertions about properties and property hierarchies. The ABox describes property assertions between individuals and membership assertions Figure 2 shows an example of these conceptual components. SPARQL is a semantic query language for semantic data, and has become a W3C standard that is able to retrieve and manipulate data stored in semantic databases (Sirin and Parsia, 2007; Harris et al., 2013; Protocol, 2016).

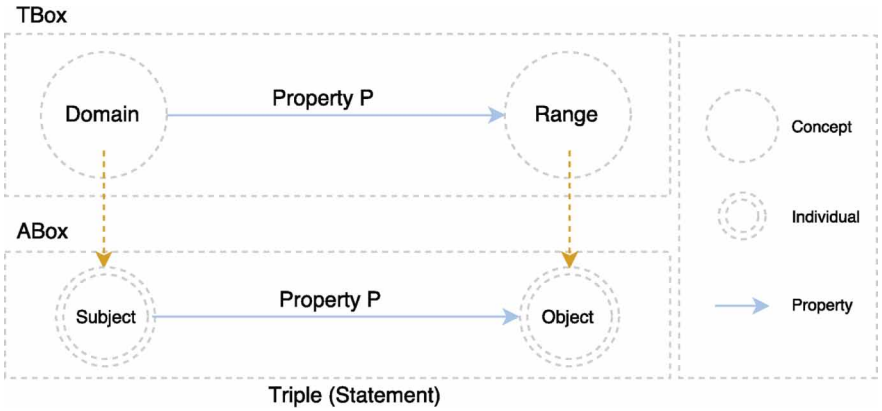
Definition 1: A DL ontology knowledge base is generally composed of (T, A) where:

- T is an intentional knowledge (TBox), which defines the concepts and properties, as well as the axioms of the logical theory;
- A is an extensional knowledge (ABox), which defines the membership of individuals (instances) and couples to concepts of individual relationships.

3. SEMANTIC HEALTH MEDIATION AND ACCESS CONTROL MANAGER

Interoperability between healthcare providers is critical for several reasons as patients may visit many physicians at different healthcare centres and their healthcare records are stored at multiple locations in disparate EHRs. The core goal of the proposed model is to enable physicians to exchange EHR data

Figure 2. An example of TBox and ABox and the relations between them



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to optimize diagnosis and treatments, while enabling patients to keep track of their health records[1]. To achieve the research goal, a semantic mediation and access control manager is proposed to facilitate care coordination and improve patient safety, and reduce duplicate testing. The proposed model architecture is shown in Figure 3. The architecture has three main components: semantic-aware ontology store, semantic access manager and schema mediation. In the following, the main components of the proposed model are described. The complete algorithm for the semantic mediation and access control manager is presented in Algorithm 1.

3.1. Semantic-Aware Ontology Store

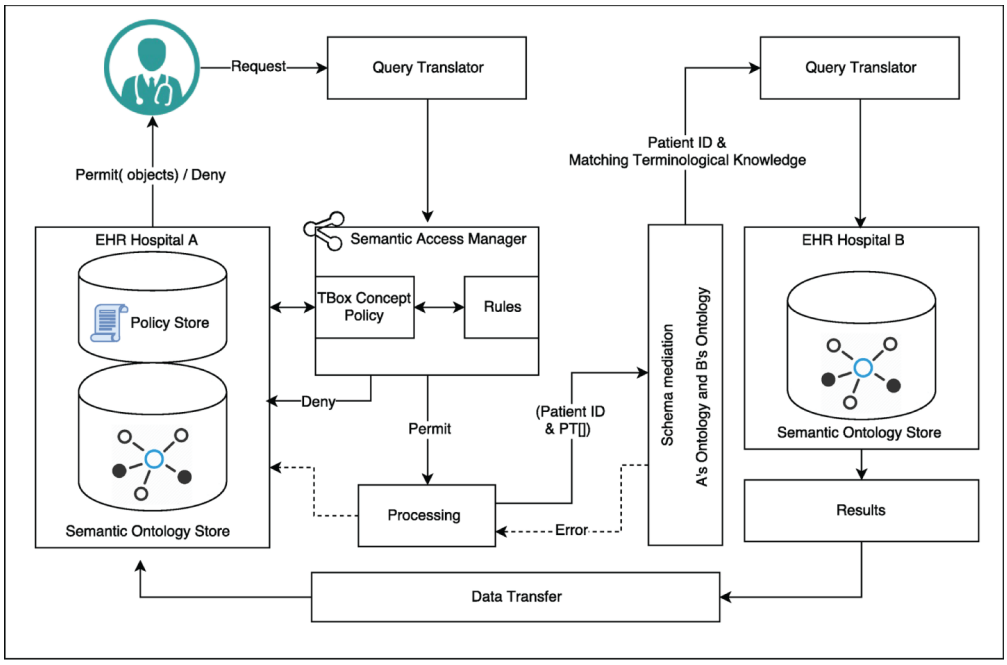
The semantic ontology stores are created for each EHR to interoperate. For this process, the knowledge base domain is represented using the Web Ontology Language (OWL), thereby obtaining the formal knowledge representation for each store.

Definition 2: The formal knowledge representation of EHR store is comprised of two typical parts:

- Terminological knowledge;
- Assertional knowledge.

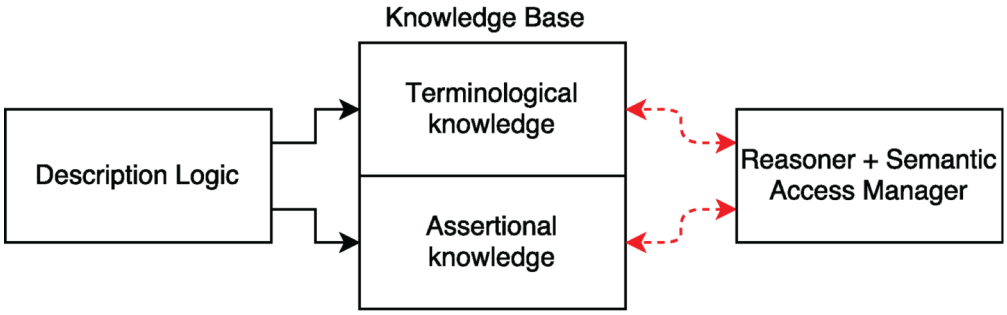
Figure 4 depicts the knowledge representation of an EHR store.

Figure 3. Semantic health mediation and access control manager



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Knowledge representation of EHR store



Definition 3: A DL ontology knowledge is defined by interpretations which is a pair $I = (D, .^I)$, where $D = \varnothing$ of individual objects, called the domain; and $.^I$ is an interpretation function which maps:

- Every concept (class) $C \sqsubseteq D$;
- Every property P is semantic relation on $D \times D$;
- Every individual x is an element $x \in D$.

In OWL, concept is an abstract group, set or collection of objects that represent the individuals (instances) in the domain of discourse. A property, which expresses relationships between two concepts in a given domain, can be further distinguished as an object property or data property. An object property describes the relation between individuals (instances) of two concepts, whereas data property describes the relation between an individual concept and the literal value.

3.2. Semantic Access Manager

Semantic access manager is specifically tailored to regulate access to the knowledge contained in semantic EHR ontology store and to protect the semantic model's concepts and their relationships based on the authorization policy. The authorizations can be expressed by the set of policy rules asserting the access privileges that a user's role has to every terminological knowledge. Therefore, the set of policy rules are used to reason over the knowledge and infer information from the knowledge based on appropriate access policies. In this section, we formalize semantic access manager with an access authorization policy. We define the role; and permission rules as the basic building blocks in the design.

Definition 4 (Syntactic Policy): The following forms the principal semantic domains underlying the semantic access manager syntactic policy:

- S is a set of subjects (users) in this policy;
- R is the role of the user in this policy;
- Pr is a set of the matching capability and permission rules;
- Let $SR \sqsubseteq S \times R$ be a subject–role assignment relation;
- Let $RPr \sqsubseteq R \times Pr$ be a role-permission assignment relation;
- Pr[i] = (Act, Tn), where Act is an operation mode and Tn is a terminological knowledge which the user is permitted to access.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

More precisely, consider two semantic ontology stores, HO(A) and HO(B), in two hospitals A & B. In Step 1, since the user does not know the semantic ontology store schema of hospital B, the user writes a query against hospital A's ontology schema about a patient. In Step 2, the query translator receives an input query. The query is parsed to identify all the target terminological knowledge that it contains. In Step 3, the semantic access manager checks its policy base for the set of access permission rules associated with the user's current role to determine whether the user/role has permission to access the target terminological knowledge in the rewritten queries. If such matching rules exist, the semantic access manager forwards a (Patient identifier, PT[]) pair to the schema mediation module. PT[] defines a set of permission terminological knowledge.

Definition 5: PT[i]= Tn indicates the authorized terminological knowledge (Tn) of the source ontology that is allowed to be accessed by the user's role.

3.3. Schema Mediation

A DL ontology knowledge has been implicitly adopted as a suitable formal knowledge representation of an EHR store. In this phase, a natural generalization is designed to formalize multiple ontology stores interconnected by semantic mappings. Ontologies correspond to description logic theories (terminological knowledge), while semantic mappings correspond to collections of a semantic mapping rule which express how subsumption migrates from semantic ontology store HO (A) to semantic ontology store HO (B) by means of the mappings. Semantic mappings between different semantic ontology stores are expressed via a semantic mapping rule. A semantic mapping rule lists two concepts from two different ontologies and the relations among them. We introduce the following definition to formalize this term.

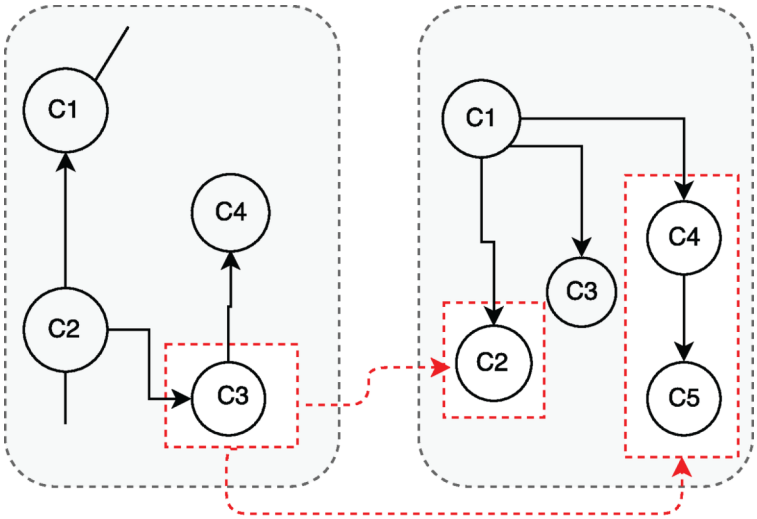
Definition 6: A semantic mapping rule is defined as: $O^A{:}C \xrightarrow{Map} O^B{:}C$.

A semantic mapping rule describes relations between HO(A) and HO(B) that are allowed to be viewed by the subject role of the A-th ontology. Therefore, the semantic mapping rule describes direct access from the source ontology to the target ontology. To make every description distinct, we will formalize each semantic ontology store with a prefix index. For example, the concept C that occurs in the A-th ontology store is denoted by $O^A{:}C$. Similarly, we use $O^B{:}C$ prefix index to refer to the concept in the B-th ontology store. The semantic mapping rule consists of three elements as stated in definition 6. Figure 5 shows an example of semantic ontology mapping:

- $O^A{:}C$: source concept, which is a concept of the source ontology that;
- $O^B{:}C$: target concept, which is a concept of the target ontology;
- $\xrightarrow{Map}$ mapping which represent the relations from the A-th ontology store O^A to B-th ontology store O^B.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

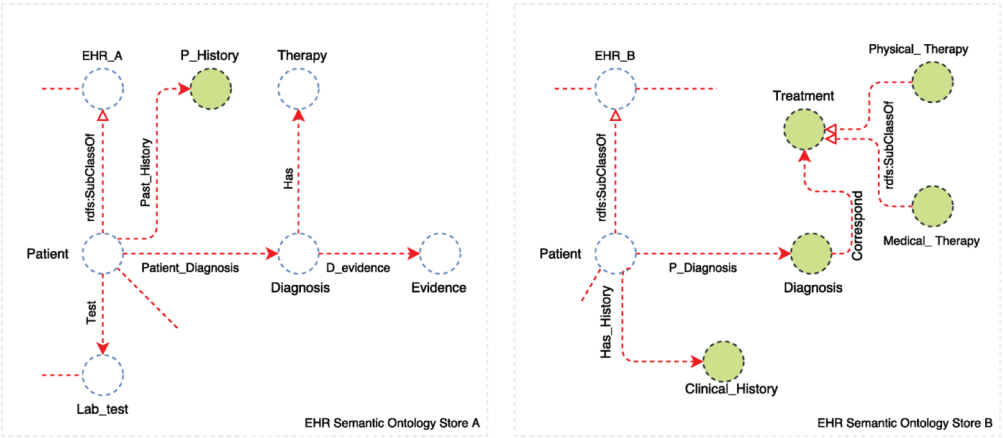
Figure 5. Example of semantic ontology mapping



4. MOTIVATIONAL EXAMPLE

The Semantic Health Mediation and Access Control Manager enables EHR data to be exchanged between healthcare organizations to optimize diagnosis and treatments, while enabling patients to keep track of their health records. The proposed model enforces access rights on a semantic terminological knowledge to secure health data within a semantic ontology store. To illustrate the procedures of the model, we create a hypothetical health knowledge model that is used to define and normalize the structures and relationships of the complex and unstructured health data from EHRs. Figure 6 illustrates a semantic ontology store for two healthcare organizations. Note that in this ontology example, we essentially focus on patients' medical records.

Figure 6. Semantic ontology store HO (A) and HO(B)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

When a user acting in a role at hospital HO(A) submits a request to retrieve the health information of a patient from semantic ontology store HO(B), the request will be processed in the following 6 steps. Step 1: As the user does not know semantic ontology store schema HO(B), the user writes the query against the HO(A) ontology schema. The query is analysed in order to identify all the target terminological knowledge that the user wants to access in the query. Let us suppose the user acting in a role in HO(A) sends a request to retrieve information about the past history of patient (ID: Patient XZ1SD). Step 2: The semantic access manager retrieves the rule policies of the user's role from HO(A) and determines whether the terminological knowledge that the user's role submitted can be accessed and then generates PT[]. Let us consider that the user's role has a rule to access the request after checking the policy base. Step 3: The semantic access manager forwards an (Patient identifier, PT[i]=C:History) pair to the schema mediation module. Step 4: Mappings between different semantic ontology stores are expressed via a semantic mapping rule. The semantic mapping rule describes relations between HO(A) and HO(B) that are allowed to be viewed by the subject role of the A-th ontology.

Algorithm 1. Semantic health mediation algorithm

Input: Semantic Query Q, Subject/Role (SR $\sqsubseteq$ S $\times$ R in HO(A))
Output: Result Objects (HO(B))
1: The Query Translator
2: Parses Q to identify all the target terminological knowledge Tn
3: Semantic Access Manager
4: Retrieves the rule policies of the user's role from HO (A)
5: Access control verify the following
6: for each Tn do
7: If SR $\xrightarrow{Permit_access}$ Tn then
8: Generate the PT[i]= Tn
9: end if
10: end for
11: Forwards an (Patient identifier, PT[]) to the Schema Mediation
12: Schema Mediation
13: Lookup the Ontology Semantic mapping rule as OA:C $\xrightarrow{Map}$ O^{B}:C
14: if not exist any mapping rule then
15: Return "Failure"
16: else
17: Retrieve the appropriate matching terminological knowledge
18: end if
19: Submit the Patient identifier with matching
20: terminological knowledge to HO(B))
21: Generate the result
22: Transfer results objects to the user's role through the mediator
23: and object transfer

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

O^A:History $\xrightarrow{Map}$ O^B:Clinical History & O^B:Diagnosis & O^B:Treatment. Step 5: The filtered authorized semantic knowledge (Patient identifier, O^B:Clinical History & O^B:Diagnosis & O^B:Treatment) are translated and executed against semantic ontology store HO(B). Finally, Step 6: The resulting objects about the patient in HO(B) are sent to the user by means of semantic mediation and object transfer.

5. CONCLUSION

In this paper, the semantic health mediation and access control manager was proposed a means of providing secure interoperation among healthcare organizations. The proposed model resolves interoperability between healthcare providers and enables access control checks which provide a secure environment. Enabling secure sharing health information among healthcare providers will enhance the quality of healthcare by reducing errors in medical decisions, decreasing health costs, and enhancing patient care. Future work will focus on implementing and improving the model algorithm (Algorithm 1).

REFERENCES

Allemang, D., & Hendler, J. (2011). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL*. Elsevier.

Ambinder, E. P. (2005). Electronic Health Records. *Journal of Oncology Practice / American Society of Clinical Oncology*, *1*(2), 57–63. doi:10.1200/jop.2005.1.2.57 PMID:20871681

Beckett, D. & McBride, B. (2004). RDF/XML Syntax Specification (Revised). *W3C Recommendation*, *10*(2.3).

Begoyan, A. (2007). *An Overview of Interoperability Standards for Electronic Health Records*. USA: Society for Design and Process Science.

Benson, T., & Grieve, G. (2016). *HL7 Version 2*. Cham: Springer International Publishing.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 28–37. doi:10.1038cientificamerican0501-34 PMID:11341160

Cannoy, S. D., & Iyer, L. (2008). *Semantic Web Standards and Ontologies in the Medical Sciences and Healthcare. In Medical Informatics: Concepts, Methodologies, Tools, and Applications: Concepts* (p. 65). Methodologies, Tools, and Applications.

El-Sappagh, S. H., El-Masri, S., Riad, A. M., & Elmogy, M. (2012). Electronic Health Record Data Model Optimized for Knowledge Discovery. *International Journal of Computational Science*, *9*(5), 329–338.

European Committee for Standardization. (2008). 251 (European Standardization of Health Informatics) ENV 13606, Electronic Health Record Communication.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Harris, S., Seaborne, A., & Prud'hommeaux, E. (2013). SPARQL 1.1 Query Language. *W3C Recommendation*, *21*(10).

Hersh, W., Totten, A., Eden, K., Devine, B., Gorman, P., Kassakian, S., Woods, S. S., Daeges, M., Pappas, M., and McDonagh, M. S. (2015). Health Information Exchange.

ISO. (2003). 215. ISO TS, 18308.

International Organization for Standardization (ISO). (2005). 2005 Health Informatics-Electronic Health Record Definition, Scope and Context Standard.

Kolias, V. D., Stoitsis, J., Golemati, S., & Nikita, K. S. (2014). Utilizing Semantic Web Technologies in Healthcare. In *Concepts and Trends in Healthcare Information Systems* (pp. 9–19). Springer.

Lezcano, L., Sicilia, M.-A., & Rodr'ıguez-Solano, C. (2011). Integrating Reasoning and Clinical Archetypes Using OWL Ontologies and SWRL Rules. *Journal of Biomedical Informatics*, *44*(2), 343–353. doi:10.1016/j.jbi.2010.11.005 PMID:21118725

Lozano-Rub'ı, R., Carrero, A. M., Balazote, P. S., & Pastor, X. (2016). OntoCR: A CEN/ISO-13606 Clinical Repository Based on Ontologies. *Journal of Biomedical Informatics*, *60*(Suppl. C), 224–233. doi:10.1016/j.jbi.2016.02.007 PMID:26911524

Mandel, J. C., Kreda, D. A., Mandl, K. D., Kohane, I. S., & Ramoni, R. B. (2016). Smart on FHIR: A Standards-Based, Interoperable Apps Platform for Electronic Health Records. *Journal of the American Medical Informatics Association*, *23*(5), 899–908. doi:10.1093/jamia/ocv189 PMID:26911829

Martinez-Costa, C., Legaz-Garc'ıa, M., Schulz, S., & Ferna'ndez-Breis, J. T. (2014). Ontology-Based Infrastructure for a Meaningful EHR Representation and Use. In *2014 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)* (pp. 535-538). IEEE. 10.1109/BHI.2014.6864420

Martínez-Villaseñor, M. L., Miralles-Pechuan, L., & González-Mendoza, M. (2016, November). Interoperability in Electronic Health Records Through the Mediation of Ubiquitous User Model. In International Conference on Ubiquitous Computing and Ambient Intelligence (pp. 191-200). Springer, Cham.

McBride, B. (2004). The Resource Description Framework (RDF) and Its Vocabulary Description Language RDFs. In Handbook on ontologies (pp. 51–65). Springer.

McGuinness, D. L., Van Harmelen, F. (2004). OWL Web Ontology Language Overview. *W3C recommendation*, *10*(10).

RDF Data Access Working Group (DAWG). (2016). RDF Query Language.

Sinha, P. K., Sunder, G., Bendale, P., Mantri, M., & Dande, A. (2012). *Electronic Health Record: Standards, Coding Systems, Frameworks, and Infrastructures*. John Wiley & Sons. doi:10.1002/9781118479612

Sirin, E., & Parsia, B. (2007). *SPARQL-DL: SPARQL Query for OWL-DL* (Vol. 258). OWLED.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Stroetman, V., Kalra, D., Lewalle, P., Rector, A., Rodrigues, J., Stroetman, K., … Zanstra, P. (2009). Semantic Interoperability for Better Health and Safer Healthcare.

Turhan, A.-Y. (2011). Description Logic Reasoning for Semantic Web Ontologies. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics*. ACM. 10.1145/1988688.1988696

ENDNOTE

[1] Note: our focus in this research is on the patient (requested data in a patient's electronic health record)

This research was previously published in the Journal of Information Technology Research (JITR), 11(4); edited by Francisco José García-Peñalvo; pages 87-98; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 10
A Conceptual Framework for an Extension Access Control Models in Saudi Arabia Healthcare Systems

Amin Shaqrah
Taibah University, Saudi Arabia

Talal Noor
Taibah University, Saudi Arabia

ABSTRACT

This article aims to develop an extension access control models framework in Saudi Arabian healthcare systems. The conceptual framework acts as an ascendency structure to organize and support the efforts of several health care standards which reflect on the coherent of confidentiality; integrity; and availability triads in order to achieve the strategic business objectives of Saudi Arabian healthcare institutions. It is considered to be three common access control models developed by ACM institute and extended to other criteria identified by the National Institute of Standards and Technology. While literature explains that an easy-to-use access control model can lead to success healthcare system, understanding the extension of access control systems is vital for Saudi Arabian healthcare institutions to protect resources against unauthorized use. This article has taken a step in this direction.

INTRODUCTION

The purpose of information security in Healthcare Information Systems (HIS), in general, is to guarantee the Confidentiality, Integrity, and Availability (CIA) of the data (Srisakthi and Shanthi, 2015). Confidentiality of the data is the protection that only those with appropriate rights and verified permissions might access certain data (Whitman et al., 2013). Data should not be disclosed to unauthorized entities, integrity in general means sustaining and ensuring the accurateness of data over its entire life cycle.

DOI: 10.4018/978-1-7998-1204-3.ch010

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In HIS, integrity of the data means that data should not be modified by unauthorized entities/persons. Protection of data in HIS prevents unofficial or accidental with holding of data or resources. To insure HIS security, countries initiate laws and regulations that healthcare organizations must follow. In the US there are three acts, the Health Insurance Portability and Accountability Act (HIPAA), the Gramm Leah Bliley Act (GLBA), and the Sarbanes Oxley Act (SOX), that involve statements to protect personal information from being revealed or retrieved by unofficial persons, entities, or processes (Gopalan et al., 2012). HIPAA upholds principles for the practice and release of Protected Health Information (PHI), which is any data about health status, provision of health care, or health care expenses that can be connected to an individual (Lerouge, et al., 2007).

To protect PHI, healthcare organizations need to enforce the patients' rights by using a set of policies and technologies. Access control models were introduced to overcome the privacy matter and grant permission to access PHI to only authorized persons. Access control models are used to prevent unofficial use of resources, including using resources in an authorized manner (Zeltsan, 2010). There are many access control models, each of these models have been extended in different ways to cover missing security measures. Healthcare organizations are free to choose the specific access control model that fits the organization's needs and is compatible with PHI privacy regulations. In addition to laws, regulations, policies and technology, standards are also used to confirm the security of the PHI in healthcare information systems (Fichman and Kemerer, 1997). Standards can be used to deliver the basic required measures to help enforcing and maintaining information security procedure in any institution. There are few well-known standardization tier-1 organizations such as the International Organization for Standardization (ISO), the National Institute of Standards and Technology (NIST), and the Association of Computing Machinery (ACM). There are other standardization organizations specific to certain industries. For example, Health Level Seven (HL7) addresses concerns about healthcare informatics. The industry standardization organizations generally adopt the technology that has been used by tier-1 standardization entities (Meyer and Goes, 1998). The literature study did not provide a methodology to extend for achieving the NISTER, the presented paper provides a starting point to meet the need for developing an extension access control model framework that addresses the synergy between secure healthcare systems and access control models considering NISTER 7874. The rest of this paper is structured as follows: Saudi Arabia health care systems, an overview of access control model, assessment access control model, proposed framework, implication, finally the conclusion and future work.

SAUDI ARABIA HEALTH CARE SYSTEMS

The unsettled environment of the healthcare sector growing the importance to adopt new technologies to achieve the vision of healthcare institutions (Samarati et al., 2001, Rossetti et al., 2012 and Hoeur and Kritchanchai, 2015). This section will provide the current requirement for healthcare systems. As mentioned in the definition of the access control, its purpose is to prevent unauthorized access to PHI. Conell and Young (2007) assertion the general privacy needs from the ACM are: 1) Healthcare institutions requirements have the ability of having the choice of designing the security policy, and the resilience to randomly define the security of PHI. 2) Patient requirements to have the ability for having full control over their PHI, right to disclose certain information, and transfer the rights from one healthcare provider to another. 3) Ease of handling ACM. Recently, the office of the National Coordinator for Health Information Technology published guidelines to preserve privacy and security related to PHI (Canada

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Health Infoway, 2015), the guidelines suggested that healthcare institutions have to address the following points: 1) auditing functions, 2) unique user IDs and strong passwords, 3) Role or User- based Access Control, 4) auto time-out, 5) amendments and accounting of disclosures (Cresswell and Sheikh, 2013). Currently, most of the healthcare systems are integrated using traditional technologies and standards (Almadani et al., 2016); security is the most important issue when integrating healthcare systems. The KSA healthcare system currently faces numerous challenges and the MOH is under remarkable pressure from its government to develop a high-quality healthcare system and improve healthcare services to all residents, mainly in rural/remote areas (Khudair, 2008). In 2010, the MOH conducted a study into the adoption of telemedicine stating that telemedicine abilities can improve many challenges of the KSA healthcare system (El-Mahalli et al., 2012). In 2011, the MOH launched the first national project for telemedicine, referred to as the Saudi Telemedicine Network (STN), covering all Healthcare Facilities (HCF) (MOH, 2016).

Alaboudi (2016) noted Saudi Arabia healthcare system exposed to various obstacles and even MOH pressures to offer high quality healthcare services to all residents, mostly in rural areas, ministry of health on (2010) initiate telemedicine adoption. Alkhamis and Hassan (2014) concluded the variances among literature with respect to adopt telemedicine. Alkhamis (2017) and Aljazaeri (2017) resolute the amount which MOH support healthcare systems. Saudi Arabian MOH deploy a complete package of health care management through the main solution to manage the "OASIS" health information system, which offers multi clinical services such as: financial, administrative and focused on facilities. The OASIS system is divided into module, each of these modules address explicit functional requirements. OASIS run on individual self-contained units, and merged with other units, and when the operating system is the integration of all units into a cohesive and powerful system. OASIS information systems in Saudi Arabia hospitals deploy different services that give the systems unique and distinctive center in the health care market, some of these features (MOH, 2016):

1. **Adjustability:** The ability to run applications in a standardized mode to suit different sizes of hospitals requirements, and these topographies and functions necessary hospital services, number of employees and specialization of the hospital play significant role. Applications are complete and very flexible and modular so that the implementation of selected modules only. This also allows the implementation of the system in stages within the hospital. OASIS system comprehensive and deals with an extensive kind of clinical systems. Simultaneously, the system is in the form of units in design, with each unit being address detailed serviceable needs. It is likely to deploy single self-contained units, and merged with other units, and when the system is implemented. All units are integrated into a cohesive and powerful system.
2. **Applicability:** The ability to run applications on a diversity of platforms diverse hardware, it is likely that the deployment of the program on different PC, unrelatedly of ability, iOS which they used starting from advanced systems in servers to small computer networks. This has significant benefits as described below. The system of the laboratory, for case, needs a lot of early group consists of selecting diverse stands such as test codes, ranging bookmark, and excellence control standards, and arrangements current work, and so on. This procedure takes a long time wants a lot of believed and effort of supervising staff. If it was likely to handover this group of raw data to another location with the application, it will save a portion of work through the provision of the laboratory application in every hospital. This would kind the work a great amount of uniformity and consistency in the hospital operations.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3. **Scalability Interdependence:** The capability to transfer hospitals on the application flat to access from everyplace inside the system, and access to the patient's health history in the past online, this to improve the excellence of treatment and decreases the total cost of patient. Scalability interdependence can be accomplished if there is an active connection on the application level through OASIS application. OASIS application delivers connectivity levels. Such a level of communication removes the stays period and imprecisions customarily related with system. And the stays period related with improved functionality many referral appointments, lab reports and at the end deliver better health care.
4. **Data Management:** The facility to retain and store data in the manner greatest favorable from the viewpoint of data management in the hospital.
5. **User Interface:** Ease of GUI browsing and navigation allows users deal quicker and nearer to PC by providing GUI directly to regulatory databases. And end users can understand the database and prepared by the system to fit their requirements. OASIS system provide a comprehensive range of applications obtainable that provides a elastic mechanism for the facades of data entry, query facilities, reports attached to the writer and GUI interface to present and analyze data visually.
6. **Arabic GUI:** Hospital system has an effect on the whole hospital, and managerial staff, paramedics and medical staff. Humiliation, the dual-language program has the capabilities to store various data about the patient and other managerial particulars in Arabic and English. OASIS after it was deployed in the Middle East proficient of supplying a perfect system. It is fantasy that health data will be in English, while the managerial data perfect. The system also allows the query and prints some of the information in any of the languages were chosen by the employer.

Furthermore, other features that may be of interest to the user like modify the forms flexibly; ease of maintenance; confidentiality and security of data; rapid response; and the availability of local expertise and support (MOH, 2016).

AN OVERVIEW OF ACCESS CONTROL MODELS

Medical information is a standout amongst the most delicate classes of data whose confidentiality and integrity is crucial. Medical information ought to then be available just to the approved social insurance experts and patients. In specific, among social insurance experts, the doctor in charge of the patient's care and medicines is generally dependable for the utilization of the patient's information (care stewardship). Patients ought to have the accessibility to get to their data and to keep up the control over who can get to it by methods for a legitimate notice system, earlier endorsement, and express assent. Since patient's wellbeing is more imperative than information classification and respectability, the standard of delivery of care starts things out must be implemented. The primary outcome of this guideline is that, if there should arise an occurrence of crises, the restorative data of a patient ought to be available to whom can convey care to the patient, regardless of the possibility that she is not unequivocally approved. In these circumstances, get to control strategies might be abrogated. To counteract or find conceivable misuse by means of a post investigation, each access conceded must be recorded in affiliation with the personality of the subject getting to the framework. This paper focused on access control model standards which they are used to provide the minimum requirements for any organization to apply its information systems security policies. In a healthcare system, these standards would be used as a minimum requirement to guarantee

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

that only authorized individuals gain access to PHI and only when needed (Barrows et al., 1996). This paper proposed a framework based on three well-known Access Control Models: Discretionary Access Control (DAC), Mandatory Access Control (MAC), and Role-Based Access Control (RBAC). DAC is one of the earliest access control models. It controls data access by authorized individuals on the basis of the user's identity and approvals that identify the access methods (i.e. read, write, or execute) that the user is permitted to view (Sandhu et l., 1994). DAC can be implemented by using a simple AC matrix or Access Control List (ACL) (Gopalan et al., 2012). DAC is a flexible policy, which makes it appropriate for different systems and applications. On the other hand, it is hard to manage when users need to be removed or added. The second well-known ACM is the MAC; it differs from DAC in the decision process. When using MAC individual data owner cannot give or modify access rights on his own data but the data is governed by a central authority not by an individual owner of the data (Sandhu and Jaehong, 2003). Additional significant field of involvement intensive on extending eXtensible Access Control Markup Language with a new security and privacy authorization policy, which adopts interoperability in the healthcare context and presents policies to impose authorization policies controlling access to information, then possibly stored across enterprise boundaries.

Since in the healthcare situation the benefits of subjects frequently rely on upon their part (e.g., nurse, doctor, patient), and get to control models to bolster both the meaning of Approvals for subjects that can be progressively composed in gatherings and the meaning of a chain of importance of parts. Moreover, get to control models and approaches for medicinal services frameworks ought to bolster the various leveled association of the data to encourage the meaning of approvals. Get to control choices may frequently rely on upon the data put away for the included patient, which incorporates distinctive sorts of information with various security necessities. As a result, a fine-grained access control that confirms regardless of whether to allow access to wellbeing data on trait/property premise is vital. The support for setting data that triggers the initiation of (dynamic) approaches is additionally of principal significance. Truth be told, an access control choice may rely on upon an arrangement of condition components that impact how and when an arrangement is upheld. For example, under ordinary conditions the wellbeing data of patients put away in the database of a healing facility can be gotten to just locally, in the occasion of a pestilence or of uncertain events, physicians working remotely can get to patient information to treat casualties. In spite of the fact that in crisis circumstances arrangements that apply in typical conditions can be abrogated, there are a few accesses that ought to dependably be avoided, since they can't help in overseeing such crisis circumstances and speak to mishandle. Such gets to ought to never be allowed, implying that strategy supersede must be denied. The access determination is made based on the confidentiality level (top secret, secret, confidential, and unclassified) of data and access class of the user. MAC is the same as DAC in the level of difficulty. Additionally, both can be difficult to manage when users need to be added or removed (Gopalan et al., 2012). The third well-known ACM is RBAC; RBAC determines access based on the user's role and the authorizations allowed by each role (Gopalan et al., 2012), RBAC supports features, like least privileges, and policy neutrality. RBAC combines flexible and mandatory access control models. It uses roles which are a normal concept for most organizations (Baracaldo and James, 2013). In RBAC, authorization is allocated to the role that individual users have within an entity (Sandhu and Jaehong, 2003). Users such as doctors, nurses, or clerks take on given roles. The rights of access for PHI are clustered by role name, and the use of resources is limited to individuals who are authorized to accept the associated role. However, the RBAC model supports permission before access and does not offer monitoring of an authorized session. So, if the doctor is accessing an authorized session, the network administrator cannot monitor the doctor activities. Also,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in RBAC updating information is based on action or canceling the session (Gopalan et al., 2012). Usage CONtrol (UCON) model was introduced by Park and Ravi that incorporates the old-fashioned ACM, digital rights supervision, and trust supervision (Park and Ravi, 2004). The previous three ACM deal only with the permission decisions on a user's access to gain access to data, But UCON was introduced as a broad view of the access control to cover obligation, continuity, mutability, and conditions (Park and Ravi, 2004). The UCON model incorporates the previous three access control models, plus, Digital Rights Management (DRM) and trust management.

NIST INTERAGENCY REPORT "7874"

In 2006 the National Institute of Standards and Technology (NIST) in the U.S. developed a guideline NISTIR-7316 for any organization that included the minimum requirements to secure its information (Hu et al., 2006). In 2012 NIST extended those guidelines with an evaluation metrics of access control NISTIR-7874 (Hu et al., 2012). In the healthcare system, Health Level Seven (HL7) is an entity that is internationally recognized for developing standards about healthcare informatics. HL7 is concerned with sharing the data between healthcare entities. HL7 adopted RBAC because of it is well-known model that had been explored by many organizations such as NIST, Association of Computing Machinery (ACM), and ISO (HIT, 2015). HL7 issued a third version of the adopted RBAC model. The catalog seeks to target Health Information Technology (HIT) merchants, HIT systems integrators, and healthcare provider organizations (Samarati et al., 2001). National Institute of Standards and Technology Interagency Report (NISTIR) used NISTIR-7316 to provide useful and theoretical guidance, to help organizations with securing their applications and delivering contextual information about access control policies, models, and mechanisms. It explains only the commonly used ACMs. It also discusses the capabilities, limitations, and qualities of these access control protocols that are set for each access control policy. It is proposed to provide useful and theoretical guidance for security administrators, managers, and procurement officers whose proficiency is associated with access control. According to Hu et al., (2006) NISTIR-7316 is divided as follows: 1) introduction, 2) overview of Access Control, 3) capabilities and limitations of access control mechanisms,4) capabilities and limitations of access control for distributed systems, 5) safety limitations, 6) and quality metric for access control systems.

For the purpose of this paper the enforcement mechanisms from NISTIR-7874 will be explained. The NISTIR-7874 explains only the commonly used ACMs which are DAC, MAC, and RBAC. It did give a brief overview, the pros, and cons on each of them. It highlighted the attention to the temporal restrictions which are formal statement of access policies that involve time-based limitations on access to resources. The temporal restrictions depend on the application scenarios. It may be required to limit resource use. Or, it may be to control time-sensitive activities (Samarati et al., 2001). The NISTIR-7874 gives an example of time restrictions statement which is the history-based access control policy. Within the temporal restriction subsection, it gives an overview on two of the popular history-based access control policies. NISTIR-7874 identifies the enforcement as the measures or methods that access control needs to impose rules of the used ACM. These properties affect the efficacy of extracting access control decisions. The enforcement properties included in the evaluation metrics guidelines are: 1) policy combination, configuration, and limitation, 2) bypass, 3) separation of duty, 4) safety, 5) conflict resolution or prevention, 6) operational/situational awareness, 7) granularity of control, 8) expression (policy/model) properties, 9) and adaptable to the implementation and evaluation of access control policies.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A CONCEPTUAL FRAMEWORK OF AN EXTENSION ACCESS CONTROL MODELS

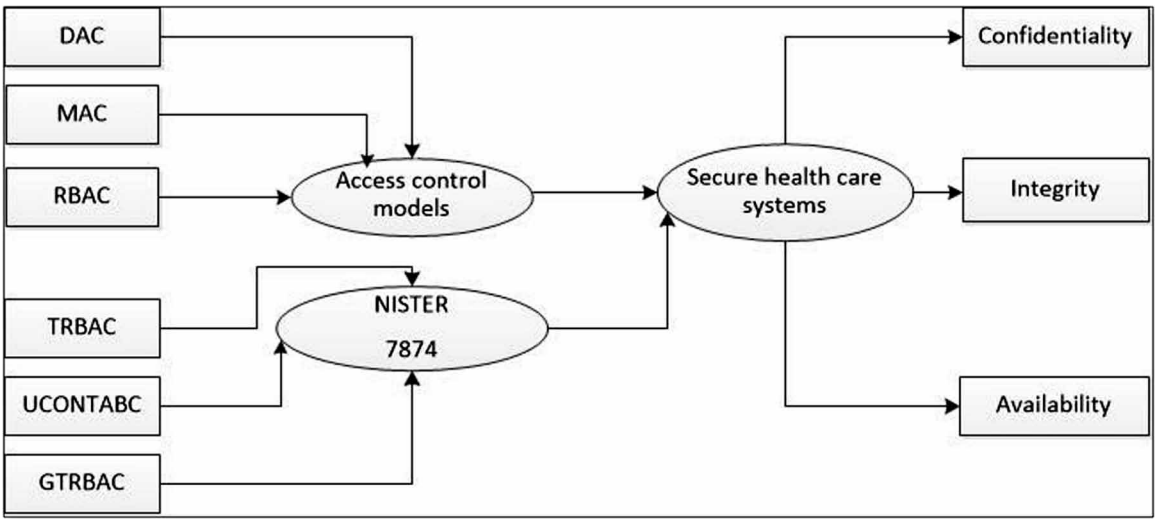
There is neither best-fit framework nor best-practice solution for telemedicine for all healthcare institution (Healy, 2008). The obstacles adopting telemedicine in a given healthcare institution may not similar telemedicine in another healthcare institution (Yusof, et al., 2008). Most healthcare institution faces expressions obstacles in adopting a specific telemedicine with an important point of dissimilarity. Nevertheless, each healthcare institution will have its own unique sets of obstacles related to its context and macro or micro environment (Yusof et al., 2008). Some of the obstacles and encounters that may limit telemedicine adoption in a given healthcare institution may no longer exist, could partially lessen, or may develop a chance for another telemedicine model healthcare institution (Baker, 2012). Perednia and Allen (1995) noted the eventual realization of adopting and implementing telemedicine in a given healthcare institution is confirmed if these obstacles and encounters are sufficiently addressed. As mentioned, each of the well-known access control model had been extended in different ways to cover missing security measures. This section will discuss some of those extensions based on our hypothetical scenario and how the researcher develops a conceptual framework overcame the missing security measures (Figure 1). The first requirement was temporary permission that should be granted by the access control system to guarantee privacy of PHI.

The temporary permission was introduced to access a control model such as Temporal-RBAC (TRBAC). TRBAC was an extension of RBAC meant to sustain time-based controls to enable/disable roles (Bertino et al., 2001). The main feature of the TRBAC model was to include periodic role permissions and time-based requirements among roles that can be identified over and done with initiators. TRBAC was extended to Generalized-TRBAC (GTRBAC) by creating a clear difference between role permission and role initiation (Joshi et al., 2002). So, the role is permitted if a user can get the authorizations given to it. A role would become active when a user gets authorization. GTRBAC was extended by hybrid role hierarchy (Joshi et al., 2005). The third requirement in the hypothetical scenario is to provide a purpose for access PHI. In an access control system, the purpose can be identified as the reasons for accessing or collecting private information. Lili et al., (2012) proposed to integrate the purpose-based usage in $UCON_{ABC}$ (Lili, et al., 2012). $UCON_{ABC}$ is an extension to UCON and the access decision is granted based on user's Authorization, Obligation, and environmental Conditions required (ABC) (Park and Ravi, 2004). $UCON_{ABC}$ allows updating features of users and data during the approved session. In addition, $UCON_{ABC}$ offers monitoring of an authorized session not like the RBAC, updating information based on action, and canceling the session (Park and Ravi, 2004).

Baracaldo and Joshi (2013) introduced another approach that was integration between risk management and the RBAC model to mitigate insider threats. It can also be applied to a third requirement of such a scenario (Baracaldo and James, 2013). Baracaldo and Joshi's approach is a good combination between preventive (ex-ante) and corrective (ex-post) mechanisms; however it needs a network manager to prevent the risk associated with each level of authorization in the system. There are recent studies that propose an integrated approach between two or more ACMs, Khan and Ken (2012) combined two ACMs (DAC and RBAC) with a context-aware AC to get the rewards of all of them. The context-aware access control would provide the compliance with the HIPAA regulation. DAC would provide the simplicity of implementations. RBAC would provide organizational role, and limited privileges.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A Conceptual framework of an extension access control models



IMPLICATIONS

Protecting and securing PHI is very important when using ACS. Any breach of such information would reveal the patient's entire healthcare history, which is undesirable (Garcia and Effken, 2013). Unfortunately, none of the three ACMs recommended by NISTIR-7874 would fulfill the general privacy requirements. Moreover, the recommendation is for additional user-based access controls, and amendments to account for disclosure of PHI. Currently, none of the recommended ACMs meet these requirements. NISTIR-7874 and its extension evaluation metrics were used to evaluate the ex-ante measures and none of the evaluation criteria considered the ex-post enforcement measures. It was expected because all the evaluation metrics were built to evaluate older ACMs that dealt only with the authorization decisions on a user's access to gain the right of use/enter data (Rath et al., 2013). By law, it is mandatory from the healthcare organizations to provide defensive measures to protect PHI. With the current standard, NISTIR-7874 is a way behind the current law requirements. The recent approaches in the ACM are the Usage-based access control model. This model incorporates the previous three access control models as well as Digital Rights Management (DRM) and trust management. $UCON_{ABC}$ was introduced eleven years ago and it matured similarly to its predecessor, ACM. $UCON_{ABC}$ allows updating features of users and data during the approved session. It does offer ex-post enforcement mechanisms as monitoring of an authorized session; updating information based on action; or canceling the session, in addition, to the ex-ante enforcement.

From the analysis of the NISTIR-7874, it was found that the standard did include the temporal restriction which is equivalent to auto time-out in the current requirements. However, it did not include the purpose for data access, which is a basic requirement. The purpose of using the PHI is important to guarantee the PHI is used as intended and prevent unnecessary access. According to Masoumzadeh and James (2008) "the purpose of using PHI was integrated the RBAC model as an extension for it". The most recent studies focused on integrated mechanisms between ex-ante and ex-post enforcements. The recent approaches are combining ex-ante measures access control models with ex-post measures

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

such as risk management (Bertino, 2005), risk assessment (Baracaldo and Joshi, 2013), intrusion detection in RBAC (Bertino, 2005). This paper recommends Saudi Arabia MOH that should be added the NISTIR-7874 to his security policy. The first section introduces $UCON_{ABC}$ as a forth recommended for ACMs, second section is TRBAC, and third section is GTRBAC. Such a model if added to the framework would mean: industrial standardization entities like HL7 would have the opportunity to recommend it as a recommended model that can be used in the healthcare organizations; the standard would be compatible with the recent law requirements, and general recent requirements. In our framework, we extended to apply TRBAC; GTRBAC; and UCONTABC considering NISTER-7874 to improve CIA triads at Saudi Arabian ministry of health. This framework represents the first step of applying all access control models into the healthcare institutions to achieve healthcare institutions vision.

CONCLUSION

This paper aims to develop a conceptual framework for the access control solution in healthcare systems at Saudi Arabia ministry of health, to realize the developed five years plan which has positioned telemedicine as a prime transformation agent and enabler. To enable this vision this paper outcomes are expected to contribute critically in achieving the strategic health objectives of the institutions, such as: care for patients; connect providers at all levels of care; improving its public image; measure the performance of healthcare delivery; and transform healthcare delivery to a consistent, world-class standard. The main access control models were introduced and illustrated through a conceptual framework, to ensuring confidentiality, integrity, and availability (CIA). Chiasson and Davidson (2004); Healy (2008) found that "IS healthcare research is the most promising opportunities to push the contextual envelope of IS research''. Explicitly, Author deploying an MAC; DAC; and RBAC for the healthcare context, which extended add TRBAC; GTRBAC; and $UCON_{ABC}$ to proposed framework introduced as recommended access control models.

Furthermore, the Industrial standardization entities as HL7 would have the opportunity to recommend it, the purpose of using the PHI is important to guarantee the PHI is used as it intends for and prevent unnecessary access. The future work would be examining one of the integrated approaches that used the ex-post measures by using the NISTIR-7874 evaluation metrics. NISTIR-7874 founded that the standard did include the temporal restriction which is equivalent to auto time-out in the current requirements. On the other hand, NISTIR-7316 did not include the purpose from data usage which could be considered a basic requirement. There are some limitations which can address for future research, the research framework needs to be tested empirically, although the access control models have been measured in previous research.

REFERENCES

Al-Jazaeri, A., Ghomraoui, F., Al-Muhanna, W., Saleem, A., Jokhadar, H., & Aljurf, T. (2017). The impact of healthcare privatization on access to surgical care: Cholecystectomy as a model. *World Journal of Surgery*, *41*(2), 394–401. doi:10.100700268-016-3739-z PMID:27734081

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Alaboudi, A., Atkins, A., Sharp, B., Balkhair, B., Alzahrani, M., & Sunbul, T. (2016). Barriers and challenges in adopting Saudi telemedicine network: The perceptions of decision makers of healthcare facilities in Saudi Arabia. *Journal of Infection and Public Health*, *9*(6), 725–733. doi:10.1016/j.jiph.2016.09.001 PMID:27649882

Alkhamis, A. (2017). Critical analysis and review of the literature on healthcare privatization and its association with access to medical care in Saudi Arabia. *Journal of Infection and Public Health*, *10*(3), 258–268. doi:10.1016/j.jiph.2017.02.014 PMID:28343793

Alkhamis, A., Hassan, A., & Cosgrove, P. (2014). Financing healthcare in Gulf Cooperation Council countries: A focus on Saudi Arabia. *The International Journal of Health Planning and Management*, *29*(1), 64–82. doi:10.1002/hpm.2213 PMID:23996348

Almadani, B., Saeed, B., & Alroubaiy, A. (2016). Healthcare systems integration using Real Time Publish Subscribe (RTPS) middleware. *Computers & Electrical Engineering*, *50*, 67–78. doi:10.1016/j.compeleceng.2015.12.009

Baker, J. The technology-organization-environment framework. (2012). Information Systems Theory, Springer, 231-245. doi:10.1007/978-1-4419-6108-2_12

Baracaldo, N., & James, J. (2013). An adaptive risk management and access control framework to mitigate insider threats. *Computers & Security*, *39*, 237–254. doi:10.1016/j.cose.2013.08.001

Barrows, J., Randolph, C., & Paul, D. (1996). Privacy, confidentiality, and electronic medical records. *Journal of the American Medical Informatics Association*, *3*(2), 139–148. doi:10.1136/jamia.1996.96236282 PMID:8653450

Bertino, E. (2005). Intrusion detection in RBAC-administered databases. In Proceedings of the 21st annual computer security applications conference. IEEE.

Bertino, E., Bonatti, P. A., & Ferrari, E. (2001). TRBAC: A temporal role-based access control model. *ACM Transactions on Information and System Security*, *4*(3), 191–233. doi:10.1145/501978.501979

Canada Health Infoway. (2015). *A telemedicine roadmap for the Kingdom of Saudi Arabia,* confidential report to Saudi Arabia.

Chiasson, M., & Davidson, E. (2004). Pushing the contextual envelope: Developing and diffusing IS theory for health information systems research. *Information and Organization*, *14*(3), 155–188. doi:10.1016/j.infoandorg.2004.02.001

Connell, D., & Young, P. (2007). Evaluating healthcare information systems through an enterprise perspective. *Information & Management*, *44*(4), 433–440. doi:10.1016/j.im.2007.04.002

Cresswell, K., & Sheikh, A. (2013). Organizational issues in the implementation and adoption of health information technology innovations: An interpretative review. *International Journal of Medical Informatics*, *82*(5), 73–86. doi:10.1016/j.ijmedinf.2012.10.007 PMID:23146626

El-Mahalli, A., El-Khafif, S., & Al-Qahtani, M. (2012). Successes and challenges in the implementation and application of telemedicine in the eastern province of Saudi Arabia. *Perspectives in Health Information Management*, *9*(1), 1–26. PMID:23209455

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fichman, R., & Kemerer, F. (1997). The assimilation of software process innovations: An organizational learning perspective. *Management Science*, *43*(10), 1345–1363. doi:10.1287/mnsc.43.10.1345

Garcia, D., & Effken, J. (2013). Development and initial evaluation of the clinical information systems success model (CISSM). *International Journal of Medical Informatics*, *82*(6), 539–552. doi:10.1016/j.ijmedinf.2013.01.011 PMID:23497819

Gopalan, R., Annie, A., & Jon, D. (2012). UCON LEGAL: a usage control model for HIPAA. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (pp. 227-236). 10.1145/2110363.2110391

Healy, J. (2008). *Implementing e-health in developing countries: guidance and principles*. ICT Application and Cyber security Division.

Hoeur, S., Kritchanchai, D. (2015). Key performance indicator framework for measuring healthcare logistics in ASEAN. In *Toward Sustainable Operations of Supply Chain and Logistics Systems, Springer chapter* (pp. 37-50).

Hu, V., David, C., Ferraiolo, D., & Richard, K. (2006). Assessment of access control systems NISTR 7316. US Department of Commerce, National Institute of Standards and Technology, 1-60.

Hu, V., Karen, C., & Ann, K. (2012). Guidelines for access control system evaluation metrics. US Department of Commerce NISTR 7874, National Institute of Standards and Technology, 1-49. doi:10.6028/NIST.IR.7874

Joshi, J., Bertino, B., Latif, E., & Ghafoor, A. (2005). A generalized temporal role-based access control model. Knowledge and Data Engineering. *IEEE Transactions*, *17*(1), 4–23.

Joshi, J., Bertino, E., & Ghafoor, A. (2002). Hybrid role hierarchy for generalized temporal role based access control model. In *Proceedings of the Computer Software and Applications Conference* (pp. 951-956). 10.1109/CMPSAC.2002.1045128

Khan, M., & Ken, S. (2012). Toward a synergy among discretionary, role-based and context-aware access control models in healthcare information technology. In *Proceedings of the World Congress on Internet Security* (pp. 66-70).

Khudair, A. (2008). Electronic health records: Saudi physicians' perspective. In *Proceedings of the 5th IET seminar on appropriate healthcare technologies for developing countries*. 10.1049/ic:20080575

Lerouge, C., Mantzana, V., & Wilson, A. (2007). Healthcare information systems research, revelations and vision. *European Journal of Information Systems*, *16*(6), 669–671. doi:10.1057/palgrave.ejis.3000712

Lili, S., Wang, H., & Soar, J. (2012). Purpose based access control for privacy protection in e-healthcare services. *Journal of Software*, *7*(11), 2443–2449.

Masoumzadeh, A., & James, J. (2008). PURBAC: Purpose-aware role-based access control. In On the Move to Meaningful Internet Systems OTM 2008 (pp. 1104-1121). Springer.

Meyer, A., & Goes, J. (1998). Organizational assimilation of innovations: A multilevel contextual analysis. *Academy of Management Journal*, *31*(4), 897–923. doi:10.2307/256344

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MOH. (2016). Achievements. retrieved from http://www.moh.gov.sa/depts/IT/Pages/achievements.aspx

Park, J., & Ravi, S. (2004). The UCON ABC usage control model. *ACM Transactions on Information and System Security*, *7*(1), 128–174. doi:10.1145/984334.984339

Perednia, D., & Allen, A. (1995). Telemedicine technology and clinical applications. *Journal of the American Medical Association*, *273*(6), 483–448. doi:10.1001/jama.1995.03520300057037 PMID:7837367

Rath, A., Thavymony, S., & Jean, N. (2013). Towards purpose enforcement model for privacy–aware usage control policy in distributed healthcare. *International Journal of Security and Networks*, *8*(2), 94–105. doi:10.1504/IJSN.2013.055943

Rossetti, M., Buyurgan, N., & Pohl, E. (2012). Handbook of Healthcare System Scheduling, Springer Science, 245–280.

Samarati, P., & de Vimercati, S. C. (2000, September). Access control: Policies, models, and mechanisms. In International School on Foundations of Security Analysis and Design (pp. 137-196). Springer, Berlin, Heidelberg.

Sandhu, R., & Jaehong, P. (2003). Usage control: A vision for next generation access control. In Computer Network Security (pp. 17-31). Springer.

Sandhu, R., Pierangela, S., & Samarati, S. (1994). Access control: Principle and practice. *IEEE Communications Magazine*, *32*(9), 40–48. doi:10.1109/35.312842

Srisakthi, S. Shanthi, A. (2015). Towards the design of a secure and fault tolerant cloud storage in a multi-cloud environment. *Information Security Journal: A Global Perspective, 24*(4), 109–117.

The office of the national coordinator for health information technology. (2015). *A Guide to Privacy and Security of Electronic Health Information*, Version 2.0, retrieved from https://www.healthit.gov/sites/default/files/pdf/privacy/privacy-and-security-guide.pdf

Whitman, M., & Herbert, M. (2013). *Management of Information Security*. Cengage Learning.

Yusof, M., Papazafeiropoulou, A., Paul, R., & Stergioulas, L. (2008). Investigating evaluation frameworks for health information systems. *International Journal of Medical Informatics*, *77*(6), 377–385. doi:10.1016/j.ijmedinf.2007.08.004 PMID:17904898

Zeltsan, Z. (2010). ITU-T Recommendation X. 805 and its application to NGN.

This research was previously published in the International Journal of Knowledge-Based Organizations (IJKBO), 8(2); edited by John Wang; pages 42-52; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 11
Strategies of Virtual Screening in Medicinal Chemistry

Giovanna Ilaria Passeri
Molecular Discovery Ltd., UK

Lydia Siragusa
Molecular Discovery Ltd., UK

Daniela Trisciuzzi
Università degli Studi di Bari "Aldo Moro", Italy

Francesco Leonetti
Università degli Studi di Bari "Aldo Moro", Italy

Domenico Alberga
Università degli Studi di Bari "Aldo Moro", Italy

Giuseppe F. Mangiatordi
Università degli Studi di Bari "Aldo Moro", Italy

Orazio Nicolotti
University of Bari, Italy

ABSTRACT

Virtual screening represents an effective computational strategy to rise-up the chances of finding new bio-active compounds by accelerating the time needed to move from an initial intuition to market. Classically, the most pursued approaches rely on ligand- and structure-based studies, the former employed when structural data information about the target is missing while the latter employed when X-ray/NMR solved or homology models are instead available for the target. The authors will focus on the most advanced techniques applied in this area. In particular, they will survey the key concepts of virtual screening by discussing how to properly select chemical libraries, how to make database curation, how to applying and- and structure-based techniques, how to wisely use post-processing methods. Emphasis will be also given to the most meaningful databases used in VS protocols. For the ease of discussion several examples will be presented.

DOI: 10.4018/978-1-7998-1204-3.ch011

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

Virtual screening (VS) is a computational technique developed since early 1990s to meet the needs of pharmaceutical companies to maximize their profit when minimizing time and cost (Lavecchia & Di Giovanni, 2013). It is well known that developing a new drug is a complex process, which can take approximately 12-15 years and huge investments of about $1 billion (Hughes et al., 2011).

In the last decades, pharmaceutical companies have radically changed the approach to the discovery and development of new drugs (Everts et al., 2017; Mangiatordi et al., 2016; Myers & Baker 2001) in the attempt to increase competitiveness and innovativeness. The need to speed-up the entire pipeline of drug discovery and development has given new credit to VS approaches as a viable alternative to the High-Throughput Screening (HTS) (Shoichet, 2004; Klebe, 2006). In this scenario, VS has gained a key role for anticipating the discovery of novel hits by reducing the likelihood of failure and, thus, time and costs (Kar& Roy, 2013).

In addition, VS is effective to investigate ligands promiscuity for unveiling polyphamacological actions for the design of multi-target compounds and/or to detect off- target affinity (Chaudhari et al., 2017). The old paradigm of drug design "one-drug one-target" has been overtaken by the need of having polypharmacological drugs enabling to replace combination therapy with multitarget therapy (Anighoro et al., 2014). VS could be also valuable to draw a preliminary toxicological spectrum and especially about drug efficacy and safety (Chaudhari et al., 2017; Nicolotti et al., 2014). In this respect, main causes of market withdrawal are associated with the occurrence of idiosyncratic adverse liver effects, which are often fatal for people (M. Chen et al., 2011; Fontana et al., 2009). Well known is the case of nimesulide, a non-steroidal anti-inflammatory drug (NSAID) marketed as one of the first selective COX-2 inhibitors and currently banned from several countries in Europe and USA.

Traditionally, VS relies on two main strategies that are ligand-based (LBVS) and structure-based (SBVS) virtual screening approaches (Lionta et al., 2014; Nicolotti et al., 2008). The LBVS strategies are based on the assumption that similar chemical structures should exhibit similar binding properties with respect to a given target (Klabunde, 2007; Koeppen et al., 2011). To carry out LBVS studies, at least one active compound (e.g., agonist for a given target) and a pool of structurally similar compounds are necessary (Koeppen et al., 2011; Rognan, 2017).

On the other side, SBVS approaches are employed when the 3D solved structure, or as alternative a reliable homology model, of a target protein is available. For the sake of clarity, homology models having at least 50% sequence identity compared to the template are suitable for SBVS, being docking studies usually able to globally reproduce the reference interactions with high level of accuracy (Bordogna et al., 2011; Cavasotto & Phatak, 2009; Hillisch et al., 2004; Oshiro et al., 2004).

Generally, a SBVS approach is based on two basic steps: (1) one or more libraries of compounds are screened by biasing the active site of the targets protein (or its model) with a number of different molecular poses; (2) the compound-target affinity is evaluated by means of a scoring function whose value is used for ranking compound. A post-processing phase can be also implemented to refine the selection of the top-ranked compounds for prioritizing experimental tests (i.e., *in vitro* and/or *in vivo* assays) (Rognan, 2017). An example of efficient application of SBVS is the identification of molecular scaffolds for developing selective ligands towards human histamine H4 receptor (hH4R). This study was conducted using a library of about 8.7 million 3D-structures screened on a homology model of the hH4R, with 255 compounds selected and successfully tested (Kiss et al., 2008).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nowadays, various workflow management tools have been developed to facilitate the automated application of VS strategies (Glaab, 2016). Widely applied are the open-source software KNIME (Beisken et al., 2013; Mazanetz et al., 2012) and the commercial Pipeline pilot (Accelrys http://accelrys.com/). Several other tools have also available including Taverna (Oinn et al., 2004), KDE Bioscience (Lu et al., 2006), Galaxy (Goecks et al., 2010), Kepler (Ludäscher et al., 2006) and SOMA2 (Lehtovuori &Nyrönen, 2006).

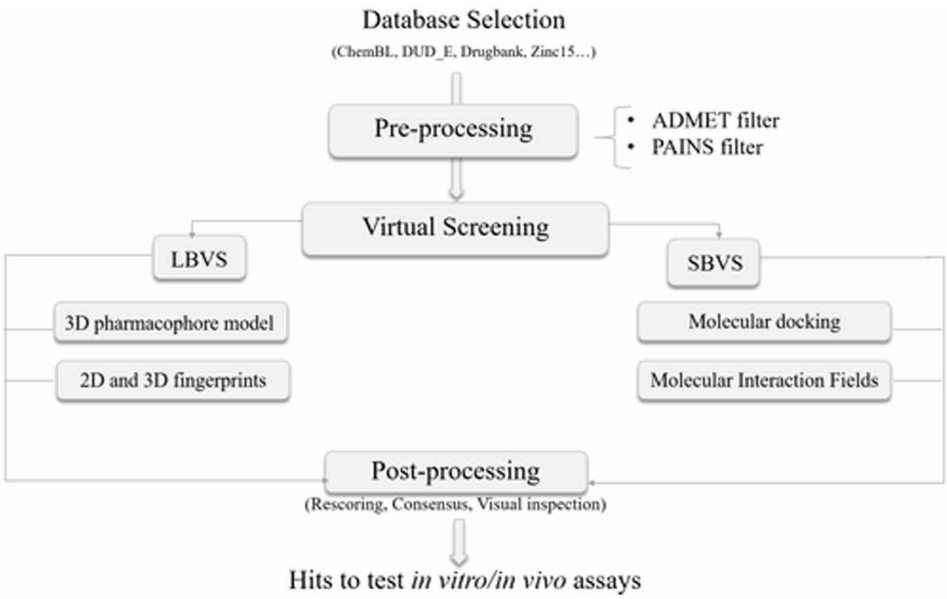
Recently, an innovative SBVS approach exploiting the similarity of X-ray solved binding sites has been developed (Siragusa et al., 2015). It is thus demonstrated that VS is also effective to explore the protein cavities whose 3D structures are known on the basis of the shape similarity and irrespective of the primary structures. Several software (e.g., BioGPS (Siragusa et al., 2015), IsoMIF Finder (Chartier et al., 2016), SuMO (Jambon et al., 2003)) able to compare the binding sites of protein structures have been so far developed.

In the first part of this review, we will survey relevant methodological steps of VS protocol (see Figure 1) including the detailed description of the two principal approaches of VS, LBVS and SBVS. In the second part, some case studies are reported in order to illustrate the pros and cons of this simulating technique.

2. LIBRARY SELECTION AND DATABASE PREPARATION

From a methodological point of view, the first step of VS approach involves the selection of a molecular library exploring a wide range of structural diversity. The selection of asuitable database is essential for increasing the rate of success in VS and computer-aided drug discovery.

Figure 1. Methodological steps of VS protocol



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Basically, molecular libraries can be distinguished as follows: 1) Focused library, library of compounds with certain biological activity (e.g., anticancer, antiviral, analgesic, antibacterial classes); 2) Targeted library, a library built against specific receptors collected into protein class (e.g., "GPCR-targeted libraries"); 3) General library, a library comprising molecules spanning different biological activities.

A number of online databases suitable for VS are known (see Table 1 for additional details). For instance, the ZINC (Irwin & Shoichet, 2005) database is a free collection of commercially-available compounds widely used for VS (Mishra & Siva Prasad, 2011; Chang et al., 2015). The latest release ZINC15 (Sterling & Irwin, 2015) contains over 100 million purchasable chemicals in 3D formats.

The National Cancer Institute (NCI) built a free database of about 250,000 structures available for searching (Ihlenfeldt et al., 2002). For each compound, an accession number (NSC) is available. For each tested chemical, the assay screen result can be indicated as active, moderately active or inactive. Specifically, the test outcomes of the Developmental Therapeutic Program (DTP) AIDS Antiviral Screen have also included.

Pubchem is an open access database structured in three linked databases, i.e., Substance, Compound (Kim et al., 2016) and BioAssay (Wang et al., 2017). These three databases are hosted by the National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM) and National Institutes of Health (NIH). To date, it contains over than 63 million unique structures and more than 1 million BioAssays collected by over 80 international organizations and scientific institutions.

Drugbank 5.0 is a freely accessible database providing a wealth of information about drugs, their targets and pharmacological actions (Law et al., 2014) resulting as a helpful drugs encyclopedia. It comprises 8261 drug entries including 2021 small molecules and 233 biotech (protein/peptide) drugs approved by the Food and Drug administration (FDA). Additionally, data information of 94 nutraceuticals and over 6,000 experimental drugs has also been collected.

ChemBL (Gaulton et al., 2012) is an open manually curated database. Currently, it counts over 2 million of2D structures and a total of 11,538 biological targets. Information concerning physico-chemical or ADMET properties for a large amount of drug-like bioactive compounds are also collected. Data are primary retrieved from literature.

DUD-E (Directory of Useful Decoys Enhanced) (Mysinger et al.,2012) database contains 22,886 clustered ligands obtained from ChemBL and their affinities against 102 biological targets. A total of 50 decoys is available for each active compound having similar physico-chemical properties but dissimilar 2D topologies taken from ZINC database. A previous reduced version, known as DUD (Directory of Useful Decoys), is also available containing 2950 ligands and 40 targets. For each ligand, 36 decoys were generated and a total of 98,266 compounds was collected (Huang et al., 2006). ChemSpider (Williams, 2011) is a free chemical structure database owned by the Royal Society of Chemistry. It provides fast text and structure search access to over 58 million structures from hundreds of data sources.

ChemDB (Chen et al., 2005) is a public database of small molecules built merging digital catalogues of over a hundred version and other public sources.

Commercial Compound Collection (Co.Co.Co.) (Del Rio et al., 2010) is a public wide repository that collects molecular structural information of almost 3.7 million of commercial compounds from various chemical vendors. Co.Co.Co. library represents drug-like and lead-like compounds in ready-to-use format to carry out VS campaigns.

In drug discovery, the keyword is innovation (Rognan, 2017). For this reason, database collecting new synthesizable molecules to explore new original scaffolds and a new possible patent space are very much attractive. The chance of identifying new hit compounds, potentially provided with a biological

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

effect, increases at the enlarging of chemical space of the library. In this respect, one should consider that the chemical space includes about 10^{60} theoretically synthesizable molecules whereas standard molecular libraries are in the range of10^6 compounds (Lavecchia & Di Giovanni, 2013). To the best of our knowledge, GDB-17 database is the largest available database with 166 billion organic small molecules (Ruddigkeit et al., 2012). It enumerates small molecules containing up to 17 atoms. A previous version GBD-13 (Blum & Reymond, 2009) counting about970 million drug-like small molecules composed of up to 13 atoms of C, N, O, S and Cl following simple chemical stability and synthetic feasibility rules are also available.

The Protein Data Bank (PDB, http://www.rcsb.org/) is a public worldwide archive containing all the information of experimentally-determined structures of proteins, nucleic acids, and complexestaken from X-ray crystallography, NMR spectroscopy or, alternatively, by homology models (Berman et al., 2000).

ModBase (https://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi) (Pieper et al., 2014) represents a collection of comparative protein structure models obtained from advanced computer-based modeling. In addition, ModBase gathers information about putative ligand binding sites, single nucleotide polymorphism (SNP) annotations and protein-protein interactions. In Table 1, all the information concerning the online chemical databases is reported. Obviously, the selection of a given database should ensure a large coverage of the chemical space according to project requirements in order to ensure efficient VS campaigns (Blum & Reymond, 2009).

3. DATA HANDLING

Data handling is the process for ensuring that data is stored and archived in a safe way in order to simplify data management as well as preserve the integrity of whole VS strategy. Two main types of data handling can be considered: 1) data curation and 2) data filtration.

Table 1. Synoptic view of free online chemical databases

Chemical libraries	URL	N. of cmpds	Last release	References
ChemBL	https://www.ebi.ac.uk/chembl	>2 million	2017	(Gaulton et al., 2012)
ChemDB	http://cdb.ics.uci.edu/	≃5 million	2005	(Chen et al., 2005)
ChemSpider	http://www.chemspider.com/	>58 million	2015	(Williams, 2011)
Co.Co.Co	http://cococo.isof.cnr.it/	≃3.7 million	2010	(Del Rio et al., 2010)
DUD	http://dud.docking.org/	≃98.000	2006	(Huang et al., 2006)
DUD-E	http://dude.docking.org/	> 22000	2012	(Mysinger et al., 2012)
Drugbank 5.0	https://www.drugbank.ca/	>8200	2017	(Law et al., 2014)
GBD-13	http://gdb.unibe.ch/downloads/	>977 million	2009	(Blum and Reymond, 2009)
GDB-17	http://gdb.unibe.ch/downloads	166 billion	2010	(Ruddigkeit et al., 2012)
NCI	https://cactus.nci.nih.gov/ncidb2.2	> 250.000	2013	(Ihlenfeldt et al., 2002)
Pubchem	https://pubchem.ncbi.nlm.nih.gov	>91 million	2017	(Y. Wang et al., 2017)
Zinc15	http://zinc15.docking.org/	>100 million	2015	(Irwin and Shoichet, 2005)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Relevant steps in data curation are summarized as follows (Mangiatordi et al., 2016; Cherkasov et al. 2014):

- Treatment of mixtures. It is a hard computational issue as several conditions are encountered: 1) the mixture containing an organic compound and several smaller moieties, either organometallic or inorganic compounds 2) several organic compounds with similar molecular weights, i.e., mixtures of geometric or stereo isomers 3) mixtures where the "inert" or impurity ingredients are unknown.
- Cleaning/ removal of counterions and salts.
- Treatment of tautomeric forms;
- Normalization of specific chemotypes. Same functional groups may be characterized by different structural patterns in the same dataset. For instance, anionic heterocycles and quaternary ammonium ions require a pre-processing and several normalization steps.
- Analysis/removal of structural duplicates. The statistical analysis of dataset requires that each compound is unique. However, structural duplicates are frequently present. If duplicates are identified, the investigation of their associated properties required a manual analysis. Duplicates must be pre-processed and removed prior to VS.
- Manual inspection. The last step of data curation involves the manual inspection of complex molecular structures.

Data filtration is frequently applied to reduce the size of database in order to remove compounds with undesirable properties. The computational filters are low-cost options to quick screen a large number of compounds (Seifert & Lang, 2008). As well known, events connected with absorption, distribution, metabolism and excretion (ADME) and chemical reactivity-related toxicity can influence drug bioavailability (Seifert & Lang, 2008). In this regard, additional filters can be applied in order to remove compounds with uncertain metabolic fate. Metal complex groups, the occurrence of large number of halogens or of unbranched alkyl chain with more than eight methylene groups can be also avoided. Other physical filters can be useful to minimize including too flexible molecules with a high number of rotatable bonds being the exploration of their conformational space too demanding (Lyne, 2002).

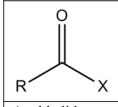
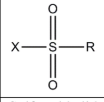
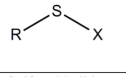
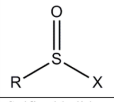
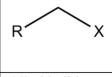
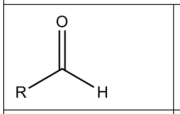
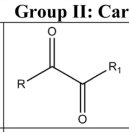
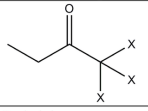
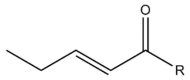
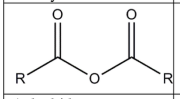
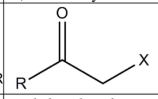
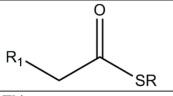
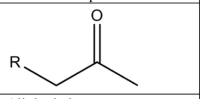
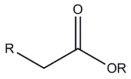
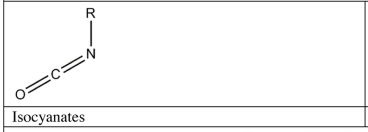
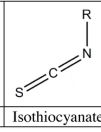
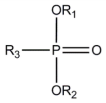
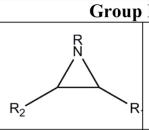
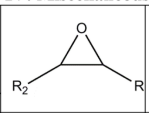
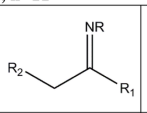
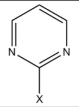
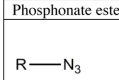
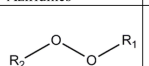
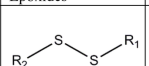
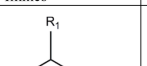
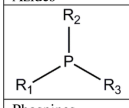
Likewise, structures with undesired functional reactive groups (i.e., Michael acceptors, ketones, aldehydes and all manner of suicide inhibitors) towards protein targets responsible of chemical instability or toxicity can be eliminated (Sirois et al., 2005; Oprea, 2002). In Table 2, all the reactive groups are listed.

There are several chemicals, known as PAINS (Pan Assay Interference Compounds), that can behave as false positive (FP) in HTS. Dahlin et al. (Dahlin et al., 2015) identified five thiol-reactive chemotypes through a series of orthogonal assay and counter-screens, with the aim to remove noise from reference database and speed up drug discovery process. In fact, as previously highlighted by Baell in 2010 (Baell & Holloway, 2010), there are features not recognizable through traditional filters that can be identified only by assay detection technology.

Moreover, compounds likely to interact with target but triggering a known adverse side effect must be also excluded. For instance, many non-antiarrhythmic drugs, characterized by different structures interact with hERG channel are responsible for a prolongation of QT interval in the electrocardiogram (long QT syndrome). This off-target interaction is associated with polymorphic ventricular tachycardia that can degenerate and cause death (Zhou et al., 1998). Thus, the filters potentially able to identify putative hERG blockers represent a valuable option to reduce cardiotoxicity optimizing VS procedure.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Reactive groups list (X = F, Cl, Br, I; R, R_1, R_2, R_3, R_4 = alkyl, aryl, heteroalkyl, heteroaryl)

Group I: Halides, n=5				
				
Acyl halides	Sulfonyl halides	Sulfenyl halides	Sulfinyl halides	Alkyl halides
Group II: Carbonyl derived compounds, n=9				
				
Aldehydes	1,2-dicarbonyl	Perhaloketones	Micheal acceptor	
				
Anhydrides	α - halocarbonyl	Thioesters	Aliphatic ketones	
				
Aliphatic esters				
Group III: Cyanates, n=2				
				
Isocyanates	Isothiocyanates			
Group IV: Miscellaneous, n=11				
				
Phosphonate ester	Aziridines	Epoxides	Imines	Halopyrimides
				
Azides	Peroxides	Disulfide	Diazo derivatives	Isonitriles
				
Phospines				

Aronov suggests that identification of hERG blockers can be achieved *in silico* through homology models, QSAR models and classification methods (Aronov, 2005).

A central issue in VS is the identification of hits compounds with a certain pharmacological or biological activity. Lipinski and colleagues (Lipinski et al., 2001) proposed simple experimental rules to establish when a molecule has drug-like properties (i.e., poor absorption or an oral bioavailability profile). These rules are summarized as follows:

- The molecular weight (MW) is less than 500Da;
- The calculated octanol/water partition coefficient (logP) is less than 5;
- No more than 5 H-bond donors – expressed as the sum of O-H and N-H groups;
- No more than 10 H-bond acceptors (expressed as the sum of N and O atoms).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Compounds matching any three of the rule of five conditions are likely to be oral drugs and can be considered as desirable chemicals with physiochemical properties that overlap with the majority of existing drugs (Lipinski et al., 2001; Cumming et al., 2013). Notably, the Lipinski's rules of five are applied as a filter for VS through a case-by-case evaluation according to the user needs.

The growing need to discover new drug-like compounds has led to develop several methods showing significant improvement of Lipinski's rules of five (Oprea et al., 2000). Based on neural network technique, Frimurer and colleagues trained a model distinguishing potentially drug-like from non-drug-like candidates (Frimurer et al., 2000). The chemicals have been taken from MACC-II Drug Data Report (MDDR) and Available Chemicals Directory (ACD). This method returns a better performance than rules of five, which accepted 74% of ACD compounds but only 66% of MDDR drugs.

Subsequently, Gozalbes and Pineda-Lucena elaborated a Quantitative Structure-Activity Relationships (QSAR) model to predict water solubility of drug-like compounds. They compared two chemical structures databases, the Food and Drug Administration Maximum Recommended Daily Dose database (FDAMMD) and Physical Properties Database (PHYSPROP) and selected the main properties (i.e., physical properties, atom count, adjacency and matrix descriptors) able to distinguish drug-like compounds. These criteria allowed to select 1174 compounds from PHYSPROP database. These chemicals have been adopted to develop several QSAR models, and the best one was selected and further validated with a set of 102 drugs whose solubility data are known in literature (Gozalbes & Pineda-Lucena, 2010).

Alternatively, bioactivity profile based on knowledge approaches can be also used for estimating the likelihood of a molecule exhibiting bioactivity. In this regard, it was reported a method to evaluate activity by analyzing what is currently known about bioactive molecules (Gillet & Nicolotti, 2000). An external chemical is likely to be active if it has physicochemical properties (e.g., such as MW, number of rotatable bonds, H-bond acceptors, H-bond donors, number of aromatic rings, shape descriptor, and ClogP) that are similar to known bioactive molecules. This approach can be also used to determine which chemicals should be screened in VS and to guide compound acquisition programms.

After Lipinski's rule, other studies were conducted. For instance, Veber (Veber et al., 2002) observed that compounds with 10 or less rotatable bonds and with polar surface area equal or less than 140 $Å^2$ showed a good oral bioavailability. Egan and coworkers (Egan et al., 2000) developed a model for human intestinal absorption using as descriptors logP and polar surface area (PSA).

Clearly, it is the combination of multiple filters that allows to prepare a database for an efficient VS. In this respect, Lagorce et al. (Lagorce et al., 2008) developed FAF-Drugs2, a free ADME/tox filtering tool useful in the drug discovery. This tool considers all the filters listed above (i.e., MW, number of rotatable or rigid bonds, heavy atoms, number of atom with a net charge, Veber's rules, Egan's rules, reactive or undesirable functional groups).

4. LIGAND-BASED VIRTUAL SCREENING (LBVS)

As above anticipated, the LB strategy exploits the principle that similar chemicals may exhibit similar properties with respect to the given target. Similarity is computed on descriptors (i.e., MW, logP, number of aromatic ring, rotatable bonds), which can be experimentally and theoretically determined (Koeppen et al., 2011).

Basically, LBVS approaches comprise QSAR strategies (Nicolotti et al., 2009; Nicolotti & Carotti, 2006; Castillo- González et al., 2015) and 3D pharmacophore models (Villoutreix et al., 2007).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4.1. QSAR Methods

QSAR is a computational technique aimed to find a correlation between a panel of chemical structures and their biological activity.

The general steps of QSAR methods are summarized as follows (Acharya et al., 2011): 1) availability of a set of congeneric molecules having a wide spectrum of biological activity as well as of chemical diversity; 2) generation of molecular descriptors accounting for structural and physico-chemical properties; 3) derivation of a causative relationships between the molecular descriptors and the biological activity; 4)internal and external validation in order to estimate the predictive power of the QSAR model.

Several methodologies can be employed to develop QSAR models such as those based on statistical approaches (Roy et al, 2015), neural networks (Aoyama et al, 1990), machine learning (King et al, 1993), random forest (Svetnik et al, 2003) and genetic algorithms (Rogers et al, 1993).

In this section, we report a brief overview of the most relevant statistical methods used to calculate the molecular descriptors as well as the main strategies employed in QSAR.

4.1.1. Statistical Methods for QSAR Model Derivation

The wider applied techniques in QSAR are the Multivariable Linear Regression analysis (MLR), Principal Component Analysis (PCA) and Partial Least Square analysis (PLS).

MLR correlates two or more explanatory molecular descriptors having a good correlation with variation in activity. However, the MLR method can be time consuming for large number of descriptors and further analyses should be carried out to avoid redundant variables. MATLAB (http://www.mathworks.com/matlabcentral) or R (Ikhara et al., 1996) are a well-known package largely employed to customize QSAR methods.

PCA analysis is often used as a diagnostic approach to extract from a large number of variables orthogonal physico-chemical information. PLS technique is instead largely adopted to build QSAR models based on latent variables largely used when using grid-like descriptors (Geladi & Kowalski, 1986). Neuronal network is the most applied non-linear regression approach to develop QSAR models (Douali et al., 2003) if no linear correlation occurs between molecular descriptor and bioactivity (Devillers, J, 2004).

Of extreme importance in QSAR is the so-called validation both internal and external. On one side, internal strategies include leave-one-out cross validation (Kohavi, 1995) and/or k-fold cross validation (Weiss & Kulikowski, 1991). On the other side, external validation is based on the predictions made on a dataset of compounds not used for QSAR model derivation (Bleeker, 2003).

4.1.2. 2D- and 3D- QSAR

Developed by Hansh-Fujita (Fujita et al., 1964), 2D-QSAR techniques are based on models where the activity contribution is determined by the substituent according to its position. This approach is effective for congeneric chemicals and is described by a simple mathematical equation including hydrophobic, electronic and steric constant. Likewise, Free and Wilson (Free & Wilson, 1964) elaborated a mathematical model in which activity depends on the presence and absence of structural elements. In particular, 2D-QSAR approaches make use of representative fragments to determine the bioactivity of a pool of compounds. However, the fragment-based descriptors are often inadequate to describe 3D conformation (González et al., 2009). 3D-QSAR includes molecular descriptors that are suitable to distinguish

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3D conformations having different importance in a spatial QSAR model (Yang, 2010, Dror et al., 2004; Reddy et al., 2007). The mainly used 3D-QSAR techniques are CoMFA, CoMSIA and Catalyst. CoMFA stands for the Comparative Molecular Field Analysis (Cramer et al., 1988) and is based on the alignment of 3D molecules. At each 3D grid point enveloping a given molecule, the value of steric and electrostatic potential energies is calculated. A limitation of the CoMFA is that typically only the lowest energy conformer is considered as bioactive molecule. Unlike CoMFA, the Comparative Molecular Similarity Indices (CoMSIA) (Klebe et al., 1994) includes hydrophobicity and hydrogen bond interactions. Several successful examples are reported in the literature. For instance, Bringman and Rummey used a dataset of 53 antimalarial biaryl compounds to develop a 3D-QSAR model based on CoMSIA descriptors (Bringmann & Rummey, 2003). In 2004, Kuo and coworkers, in a study related to HIV-1 integrase, applied a combined approach of CoMSIA and CoMFA to train a set of 66 compounds (Kuo et al., 2004). Catalyst (Kirchmair et al., 2005) is a 3D-QSAR software that consider ligand flexibility making use of poling algorithm.

4.2. Pharmacophore Models

Pharmacophore is classically defined as the 3D arrangement of molecular features of fragments required for binding then translated into descriptors (Qing et al., 2014). Recently, Ghaffar at al. (Ghaffar et al., 2017) generated a pharmacophore model to identify novel AMPK activators. Indeed, starting from a set of known active chemicals (drugs, natural compounds and xenobiotics) a pharmacophore model was generated to screen a library of ZINC database. The VS resulted hits compounds having binding energies better than the ligand itself. For instance, LigandScout (Wolber& Langer, 2005) is a software performing both LBVS and SBVS based on a 3D pharmacophore model.

Noureen et al. derived a pharmacophore model using Ligand Scout v2.02 to identify inhibitors of histone deacetylases, a key target in cancer therapy (Noureen et al., 2010). This model has been generated analysing a set of 70 compounds of three different classes. Moreover, the software has been successfully used to predict the bioactivity of novel HIV inhibitors. In this context, Barreca and coworkers identify six potential inhibitor exhibiting activity comparable to nevirapine (Barreca et al., 2007).

Alternatively, 2D and 3D fingerprints are also employed in VS to calculate properties based on the molecular structure (Baroni et al., 2007; Floris et al., 2014). Chaac (www.mindthebyte.com/saas-platform/chaac/) is a commercial software to compare a given molecular query with a database of compounds using 2D fingerprints. LiSiCA (Ligand Similarity using Clique Algorithm) (Lešnik et al., 2015) is a recent tool searching for 2D and 3D similarities between a chemical and a benchmark database. LiSiCa Plugin is free only for academic use. ACPC (Auto Correlation of Partial Charges) is an open source tool based on autocorrelation of partial charges that uses a rotation-translation invariant molecular descriptor (Berenger et al., 2014).

5. STRUCTURE-BASED VIRTUAL SCREENING (SBVS)

The number of the scientific publications reporting of SBVS approaches is approximately three times compared to those related to LBVS (Ripphausen et al., 2011). SBVS is based on the prediction of both binding energy and posing of a ligand-protein complex (Lionta et al., 2014). In this part of the review, we will focus mostly on molecular docking, a well-known *in silico* technique simulating how a small

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

molecule can interact with a target protein. Sequentially, the post-docking processes and the criteria for the selection of docked compounds for additional experimental testing will be also described.

5.1. Posing Prediction: The Ligand Flexibility Problem

Ligand bound conformation may differ from the unbounded ligand. For this reason, it is crucial to consider conformational flexibility during docking processes. Theoretically, several mathematical algorithms can simulate the molecular complementarity between the binding poses of ligandand its biological counterpart.

To correctly predict the pose, different strategies can be pursued. For instance, a database containing several conformers for each ligand can be generated and then each conformer rigidly docked on the binding site. This strategy has been implemented in several software, such as Omega (Hawkins et al., 2010) and Catalyst (Kirchmair et al., 2005).

Among others, four methods allowing ligand flexibility has been developed as follows (Sliwoski et al., 2014)

1. Incremental growth approach. In this method, a ligand is fragmented into small parts and then is re-built step by step into the binding site. This method is used, for instance, by FlexX-Scan (Schellhammer & Rarey, 2004) and Dock (Onodera et al., 2007) software.
2. Genetic algorithm. This approach searches minimum energy through analogy with genetics and evolution. In genetic analogy, the state of the ligand corresponds to the chromosome and the variables to a gene while the fitness is the interaction energy between ligand and target (Goldberg, 1989). The automated ligand docking program GOLD (Genetic Optimisation for Ligand Docking) (Jones et al., 1997) makes use of a genetic algorithm to explore a wide range of ligand conformational flexibility.
3. Exhaustive search approach (Learning, 2014). A comprehensive list of rotamer states for each rotatable bound is coupled to a heuristic screen that rapidly eliminates those conformers unsuitable for binding to the receptor. This method has been implemented in several docking programs, i.e., GLIDE (Grid-based LIgand Docking with Energetics) software (Friesner et al., 2004) available in the Schrodinger Suite.
4. Monte-Carlo methods (Metropolis et al., 1953) start from an arbitrary ligand configuration while new configurations are randomly generated within prefixed threshold (i.e., maximal changes in translational, rotational, number of torsional degrees of freedom). AutoDock (Morris et al., 1998; Huey et al., 2007) is a well-known software based on Monte Carlo approach.

5.2. Protein Target Flexibility: Is There a Way Out?

An important issue of docking-based VS is the flexibility of the target (Hou et al., 2015). Nevertheless, several efforts are also made to identify new methods that consider both the ligand and the target protein as flexible. One of these attempts is the so-called “soft docking”: the protein is still treated as a rigid structure but the criteria for steric fit between ligand and receptor are softened. For instance, Ferrari et al. (Ferrari et al., 2004) generated a “soft” scoring function weakening the Lennard-Jones potential term to screen an Available Chemical Directory (ACD) against T4 lysozyme binding site. The study resulted that soft docking shows a better performance than the classical scoring function if a single receptor conformation is used.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Other methods involved the development of algorithms that gaze at protein flexibility limited to low degrees of freedom of side-chain for selected amino acids in the active site (Flick et al., 2012). According to a real biological system, the ideal strategy would be to screen all the possible conformations of a ligand against all freedom degrees of protein structure. To bypass this computational time-demanding system, a recent approach, namely "ensemble docking", is implemented in the several docking programs considering multiple conformations of target protein (Hou et al., 2015; Xiong et al., 2015). A successful application of ensemble docking VS is reported by Hou et al. (Hou et al., 2015) that led to the identification of eight novel lymphoid-specific tyrosine phosphatase inhibitors among 23 hits compounds.

5.3. How can the Ligand-Protein Complex be Evaluated?

Typically, protein-ligand docking simulation consists of three main phases (Kitchen et al., 2004): posing, that is the process that determines the alignment and orientation of the ligand in the active site; scoring, that is the process to calculate the binding energy according to the ligand fit in the active site; ranking, that is the process to reconsider the results from previous step on the basis of a more informed and consensus view.

Scoring functions are used to calculate the strength of the non-covalent interaction between the docked compound and the target (Kitchen et al., 2004). Basically, there are three types of scoring functions developed to estimate the binding energy of a protein-ligand complex: force-field (FF), empirical and knowledge based scoring function.

FF-based scoring functions are generally based on the sum of two contributions: the ligand energy interactions and the internal energy. The former is mostly described by two terms: Van der Waals and electrostatic energy term; the latter is approximated by harmonic springs describing vibration and rotations of the different bonds forming the ligands. For instance, GOLD (Verdonk et al., 2003; Jones et al., 1997) and AutoDOCK (Huey et al., 2007) make use of FF-based scoring functions.

Empirical scoring functions are based on the idea that the experimental data can be reproduced and binding energies can be approximated to the sum of single and unrelated terms (R. Wang et al., 2002). Empirical scoring functions are preferred to shortly screen millions of compounds and a large amount of empirical data is available to train scoring function (Seifert, 2009). LUDI (Böhm, 1994), SCORE (Tao & Lai, 2001) and X-SCORE (R. Wang et al., 2002) are based on empirical methods for binding energy estimation.

Knowledge-based scoring functions are instead aimed to reproduce experimental structure on the basis of the binding geometry of known ligands. DrugScore (Gohlke et al., 2000) and Smog (DeWitte & Shakhnovich, 1996) estimate a free energy using knowledge-based scoring functions.

5.4. A New Challenge in SB Approach: Cavities/ Cavities Comparison

Recently, an increasing interest has been focused on protein binding site comparisons (Najmanovich, 2017; Chaudhari et al., 2017). In this respect, two methods for detecting binding sites have been developed: the first maps and classifies binding site pockets irrespective of any specific ligand; the second is instead ligand oriented (Lavecchia & Cerchia, 2016). In this respect, a popular algorithm for comparing binding sites is FLAP. This algorithm has been implemented in BioGPS (Siragusa et al., 2015). BioGPS is structured in 5 phases: (1) protein refinement (the Fixpdb tool); (2) cavity detection (the Flapsite tool); (3) cavity characterization; (4) cavities comparison; (5) data analysis and interpretation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The BioGPS approach has been successfully validated confirming the side effects caused by estrogen receptor modulators (Siragusa et al., 2015). Indeed, the authors demonstrated that the estrogenic compounds cause cardiac abnormalities due to an off-target effect on sarcoplasmic reticulum Ca^{2+} ion channel ATPase (SERCA).

In the last years, the importance of comparing cavities have gained breeding ground and several bioinformatics algorithms able to align the binding site sequences have been developed. For instance, the SMAP web service (Ren et al., 2010) made use of SOIPPA (Sequence Order-Independent Profile–Profile Alignment) (Xie& Bourne, 2008), an algorithm able to detect similarity pockets across all fold protein.

The sequence order of amino acid residues in the functional site is not necessarily conserved and this allows the detection of new evolutionary relationships across apparently different proteins. This software has been applied to identify off-targets for selective estrogen receptor modulators (Xie & Bourne, 2008). SiteHopper is a tool developed by Batista and colleagues (Batista et al., 2014) based on alignments of protein sequences. A 3D shape representation of the active site is evaluated and physico-chemical properties of residues are highlighted by different colors. The active site representations can be rapidly super imposed and investigated to identify shape and chemistry similarity. 3D-surfer is a free web tool widely used to compare proteins based on the surface shape (La et al., 2009). Nevertheless, it is possible to screen a protein cavity represented as a vector of 3D Zernike descriptors against the entire PDB pockets and visualize the superimposition using VisGrid and LIGSITE (Xiong et al., 2014). However, it should be mentioned that despite the increasing amount of available computational approaches developed for comparison of protein binding sites, the number of successful outcomes is still relatively low (Jalencas & Mestres, 2013) this being likely due to the difficulty of explicitly accounting in comparison for protein flexibility and structural waters. In Table 3, all the SBVS and LBVS software and related URL website are summarized.

5.5. Are the Classic Scoring Functions Now Overtaken?

The scoring functions implemented in docking programs work by simplifying the complexity of molecular events behind protein-ligand recognition. For instance, entropy and solvation energy are not taken into account. Despite the success so far obtained (Bauer et al., 2013; Odolczyk et al., 2013), new strategies have been recently developed adopting machine learning based scoring functions (Ain et al., 2015). In classical machine learning, the algorithm is trained on a set of compounds with known experimental data (Wójcikowski et al., 2017). These scoring functions are used for the binary classification (i.e., active/inactive) of external dataset of chemicals and the best performing is the linear kernel algorithm (Melville et al., 2009).

In addition to machine learning, other strategies have been proposed. A recent work (Crunkhorn, 2017) presented a new SBVS approach based on evaluation of structural stability of protein-ligand complex. Likewise, Ruiz-Carmona et al. (Ruiz-Carmona et al., 2014) described a new computational method, known as Duck, to overcome the problem of free binding energy calculation. This method was applied to the VS against molecular chaperone and target Hsp90, with promising results (Crunkhorn, 2017).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. A synoptic view of LBVS and SBVS software and related URL website

Software	URL
LBVS	
LigandScout[a]	http://www.inteligand.com/ligandscout
LiSiCA[b]	www.insilab.org/lisica/
ACPC	www.github.com/UnixJunkie/ACPC/blob/master/README.md
SBVS	
GOLD[a]	https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold//
GLIDE[a]	https://www.schrodinger.com/glide
AUTODOCK	http://autodock.scripps.edu/
BioGPS[a]	http://www.moldiscovery.com/
FLAP[a]	http://www.moldiscovery.com/
SiteHopper[a]	https://www.eyesopen.com/sitehopper
3D-SURFER	http://dragon.bio.purdue.edu/3d-surfer/
SMAP	www.smap.nbcr.net
MCDOCK	https://github.com/andersx/mcdock
Omega	https://www.eyesopen.com/
Catalyst	http://accelrys.com/
DOCK [b]	http://dock.compbio.ucsf.edu/
FlexX [a]	https://www.biosolveit.de/FlexX/

[a]commercially available
[b]an educational version is available

6. POST PROCESSING METHODS: HOW DO YOU REMOVE FALSE POSITIVES?

A limit of VS protocol is the selection of suitable docked compounds to address additional experimental tests. Keeping as low as possible the occurrence of FPs is of utmost importance in drug discovery. In this respect, different strategies of post-processing have been developed.

6.1. Efficient Post-Processing

Rescoring hits on the basis of the binding free energies (ΔG) calculation is widely adopted (Zhong et al., 2010; Kitamura et al., 2014). ΔG can be computed by Molecular Dynamics (MD) simulations or can be expressed in three energetic terms as described by Zhong et al. (Zhong et al., 2010). The authors explain that the variation in binding free energy can be calculated in vacuum through quantum mechanical (QM) Hamiltonian or molecular mechanics (MM) force fields. Alternately, it can be measured by means of solvation free energy, computed by continuum solvation models. All these systems contain an apolar component evaluated by solvent-accessible surface area (SA) models. SA models are mostly associated with the Poisson- Boltzmann (PB) and Generalized Born (GB) models referring to MM-PBSA and MM-GBSA methods, respectively. The difference in energy between the complex target-candidate ligand is also evaluated by entropic contribution. Considering ΔG as the sum of two different terms (ΔG

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$= \Delta H - T\Delta S$), high affinity is reached when both enthalpy (ΔH) and entropy (ΔS) contribute favourably (Freire, 2008). Notably, the enthalpic contribution is more difficult to optimize that the entropic one. Each attempt to improve binding enthalpy had as result a loss of entropy. For such reason, entropic contribution is preferred to evaluate the chance between target bound and unbound. As reported by Malmstrom and Watowich, another rescoring method is the mean-field pathway decoupling free energy of binding calculations (MF-FEB) (Malmstrom & Watowich, 2011). MF-FEB is a widely applied *in silico* screening even if these calculations are computational demanding (Rognan, 2017). Ding et al. (Ding et al., 2013) proposed an integrated approach called model system Molecular Interaction Energy Components - Support Vector Machine (MIEC-SVM) combining the molecular modelling and statistical learning. In this model, the molecular interaction energy (MIECs) is used as descriptor to characterize the ligand–protein interaction. The ligands are thus classified into binding and nonbinding categories training a SVM on these MIECs. The effectiveness of this model was successfully demonstrated through the screening on the HIV-1 protease using a small library of compounds (Ding et al., 2013).

6.2. Consensus Scoring

A computational strategy able to reduce FP hits is the consensus scoring (Anighoro et al., 2014). In order to decrease the scoring function-related error, a well-known strategy is to combine the scores obtained by different algorithm functions (Wang et al., 2003; Wang &Wang, 2001). It was observed that the majority of docking programs show an accuracy of about50% whereas the combination of several docking algorithms using a consensus strategy (three or four are sufficient) increases the performance of about 80% (Rognan, 2017). This method is applied using the same chemical library with different docking programs. The docking scores obtained are compared with the experimental data taken from crystallography pose and the root mean square deviation (RMSD) values are thus calculated (Houston &Walkinshaw, 2013). Similarly, the score associated to each docking pose generated by different algorithms are compared. Wang et al. (Wang & Wang, 2001), proposed three different consensus scoring strategies: these are rank by number; rank by rank, and rank by vote. In the rank by number, the candidate compound is ranked according to the average predicted value given by all the scoring functions considered. In the rank by rank, all the candidates are ranked by the average ranks predicted by all the involved scoring functions. In the rank by vote, if a candidate is predicted to be at the top list from an algorithm (e.g., in the first 10%), a "vote" from that scoring function is given so that each candidate compound will have a final vote given by the sum of votes obtained from all the scoring functions, ranging from 0 to the total number of scoring functions considered.

6.3. SBVS or LBVS?

In the decades, several studies demonstrated the success of the two mentioned strategies of VS. For instance, Xing et al., in 2011 discovered a soluble epoxide hydrolase applying the SBVS approach. The authors, starting from a library of benzoxazole, found out a new potent inhibitor for cardiovascular disease treatment (Xing et al., 2011). Spadaro and co-workers (Spadaro et al., 2012) applied a pharmacophore model to a small library of compounds active on 17β-hydroxysteroid dehydrogenase type 2 receptor. The screen identified interesting active compounds that was further chemically modified. As result, a novel class of 17β-HSD1 inhibitors was discovered.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Although the application of VS approaches is an interesting support for drug discovery studies, some necessary considerations should be done. First of all, an intrinsic limitation of LB technique is the use of starting dataset of congeneric molecules normally provided with low structural variability. SBVS overcomes this issue, benefiting from X-ray/NMR solved crystal structures. On the other side, if ligands are handled with full flexibility, the flexibility of protein target is still an open issue difficult to approach. The inclusion of structural functional water molecules in docking process could increase the accuracy of simulation (McInnes, 2007). To the best of our knowledge, SBVS is also three times more popular than LBVS strategies (Ripphausen et al., 2011).

Building on these considerations, the integrated use of structure and ligand-based approaches is, whenever possible (i.e., availability of both X-ray solved target protein and known active ligands), by great far desirable. In this respect, the interested reader can have a look at the recent study (Mangiatordi et al., 2017) relied on the hierarchical application of multilayer VS employing filters implying: molecular shape similarity, a structure-based pharmacophore model and molecular docking campaigns.

7. EXAMPLES FROM CASE STUDIES

In this section, we reported some recent case studies of VS applications showing how the VS is an effective and versatile technique for toxicological and drug discovery purposes.

7.1. LBVS

Ambure et al (Ambure et al., 2014) recently applied 3D-QSAR pharmacophore modelling in order to identify potential acetylcholinesterase (AChE) inhibitors. In particular, the InterBioScreen's Natural Compound (NC) database (InterBioScreen's NC, 2012) containing 47,868 natural compounds was screened. More specifically, a multi-filter strategy was applied involving: 1) the application of drug-like ADMET properties computed using QikProp (QikProp, 2012); 2) employment of a pharmacophore model as a 3D structural query and 3) molecular docking simulations in order to investigate the ligand-protein interactions. The 3D-Pharmacophore model was generated starting from a small dataset of 404 AChE inhibitors whose IC_{50} bioactivity is well-known in the literature. 85 chemicals, randomly selected (IC_{50} ranging from 0.950 nM to 11,587.77 nM) were adopted as training set and the remaining 319 as test set. The harmonized use of an ADMET based filter, a 3D-QSAR model and molecular docking resulted in the discovery of novel AChE inhibitors showing unreported molecular scaffolds. Hence, this study demonstrates how the wise application of LBVS allows efficiently exploring chemical space in drug discovery.

7.2. SBVS

7.2.1. Docking-based Classification Models to Predict the Estrogenic Potential of Chemicals

Trisciuzziand co-authors (Trisciuzzi et al., 2015) have established an easy-to-run platform to develop classification models based on molecular docking approach in order to predict estrogenic activity of wide collection of chemicals. The estrogen receptor (ER) is one of the most extensively studied targets (Barton et al.,2017; Almeida et al., 2017) for its role in various pathologies, such as osteoporosis (Martinaityte

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

et al., 2017), breast cancer (Gruvberger et al., 2001) and obesity (Cohen 2008; Guillaume et al., 2017). In this regard, numerous *in vitro* and *in vivo* protocols have been applied to identify potential endocrine pathway-mediated effects of well-known Endocrine Disruptor chemicals (EDCs). Nevertheless, the emergence of concerns about EDCs has resulted in the development of *in silico* strategies quickly and low cost. The presented study is framed within the large-scale modeling project CERAPP (Collaborative Estrogen Receptor Activity Prediction Project) (Mansouri et al., 2016) supervised by the US Environmental Protection Agency (US EPA) in partnership with 17 research groups in the Europe and USA.

In this work, starting from a training set (TS) of 1677 chemicals provided from US EPA, 24docking-based classification models were derived and optimized. The model performances were evaluated considering the early Enrichment Factor at 1% (EF1%) and the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. In order to classify the estrogenic potential of chemicals, three probability binding classes were defined. Notably, two thresholds were set on the basis of the tunable sensitivity (SE) values, where SE estimates the proportion of TPs that are correctly predicted at a specific position of the ranked database. The following classes were obtained: 1) $SE \leq 0.25$, the class with high probability of binding (i.e., binder toxic molecules); 2) $SE > 0.75$, the class with low probability of binding (i.e., non-binder safe molecules); 3) $0.25 < SE \leq 0.75$, the class with medium probability of binding (i.e., suspicious molecules).

For toxicological purposes, keeping as low as possible the occurrence of false negative (FNs) is of utmost relevance. In this regard, the Negative Predictive Values (NPVs) at two different SE thresholds ($SE=0.25$ and $SE=0.75$) were also computed. In addition, to overlap the intrinsic limitation of unbalanced starting TS (only 14.13%chemicals were designated as estrogenic chemicals), the positive (at $SE = 0.25$) and the negative (at $SE = 0.75$) likelihood ratio for each of the SE based thresholds have also calculated. A small dataset of 10 well-known compounds (i.e., 5 drugs and 5 widely employed chemicals) and an external blind set shared by the US EPA of about 32,000 compounds have been thus screened against the best performing classification model. To overcome the limitation of single models provided by each CERAPP participants, a consensus strategy has been implemented. The success of this innovative application is confirmed placing as fourth better model of 21 models obtained by the CERAPP initiative. The use of molecular docking, a SB strategy widely applied in drug discovery program, is now adapted to toxicological purpose.

7.2.2. Challenging AQP4 Druggability for NMO-IgG Antibody Binding Using Molecular Interaction Fields and Molecular Docking Approach

The aquaporins are water channels with a physiological role in several organs and tissues. The alteration determining a variation in the rate of water flux through the channel is associated with several disorders. Among the13 isoforms of the aquaporin, the human aquaporin 4 (hAQP4) is also well-documented for its implication in various pathologies including Neuromyelities Optica (NMO). In this regard, a deposition of a typical immunoglobulin, namely NMO-IgG, against isoform hAQP4 was observed in many NMO patients (Wingerchuk et al. 2007).

A recent study (Pisani et al., 2014) has demonstrated that residue D69 at the basis of the loop A plays a key role for AQP4-IgG binding. Hence, developing drugs able to block the interaction between hAQP4 and NMO-IgG at this level might be an interesting therapeutic approach.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Keeping this in mind, Mangiatordi et al. (Mangiatordi et al., 2015; Mangiatordi et al., 2016) investigated the loop A of hAQP4 by means of generating molecular interaction fields (MIF) on the X-crystal structure of hAQP4 (PDB code: 3GD8) using BioGPS software. The sole cavity identified by the FLAPsite algorithm was located at the tophAQP4 surface confirming the previous hypothesis about the relevance of loop A. In order to demonstrate the druggability of this region, a successful analysis was performed against a benchmark dataset of druggable cavities of 43 X-ray solved pharmaceutically targets. In particular, the computed hydrophobic, hydrophilic, and charge-charge/polar interaction descriptors of the cavity fell in the range of those calculated for the druggable dataset confirming the potential druggability of hAQP4.

Moreover, according to the recent experimental observations, an induced-fit docking simulation of three known drugs (i.e., arbidol, tamarixetin and berbamine) having an appreciable inhibition of hAQP4NMO-IgG binding was thus performed. Satisfactorily, all the three compounds were found and almost completely inside the candidate hAQP4 cavity at the top of loop A.

8. CONCLUSION

Although turning its 30th birthday, VS is still a very effective method largely applied in medicinal chemistry. To the best of our knowledge, we illustrate the main methods currently employed for VS presenting some suggestive case studies to guide beginners in this new fascinating field. Moreover, a large room was given to survey molecular databases as they represent the wealth of information that is critical for VS success. Far from being an exhaustive analysis about VS, this review represents a comprehensive and informed view of the most applied strategy mainly in medicinal chemistry program. Indeed, much has been already done in this area but still a lot is in front of us. An interesting future perspective is to integrate as much as possible LB and SB strategies. In this respect, VS can be critical for repurposing old drugs towards new targets, for better designing multi-target agents and for addressing in a next future the large-scale screening of antibodies, whose clinical role is central for personalized medicine protocol. In this respect, finding new and cheaper solutions to approach with protein flexibility is of utmost importance to gather more reliable and trustable results even provided of a solid mechanistic rationale. Finally, we believe that VS could represent a valid starting point to find a hit for the 'hit to lead' approach contributing significantly to the whole drug discovery process. Nevertheless, the visual inspection represents the final filter whose success is very much dependent by chemical feeling and user intuition.

REFERENCES

Acharya, C., Coop, A., Polli, J. E., & Mackerell, A. D. (2011). Recent advances in ligand-based drug design: Relevance and utility of the conformationally sampled pharmacophore approach. *Current Computer-aided Drug Design*, *7*(1), 10–22. doi:10.2174/157340911793743547 PMID:20807187

Ain, Q. U., Aleksandrova, A., Roessler, F. D., & Ballester, P. J. (2015). Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *5*(6), 405–424. doi:10.1002/wcms.1225 PMID:27110292

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Almeida, M., Laurent, M. R., Dubois, V., Claessens, F., OBrien, C. A., Bouillon, R., ... Manolagas, S. C. (2017). Estrogens and Androgens in Skeletal Physiology and Pathophysiology. *Physiological Reviews*, *97*(1), 135–187. doi:10.1152/physrev.00033.2015 PMID:27807202

Ambure, P., Kar, S., & Roy, K. (2014). Pharmacophore mapping-based virtual screening followed by molecular docking studies in search of potential acetylcholinesterase inhibitors as anti-Alzheimers agents. *Bio Systems*, *116*, 10–20. doi:10.1016/j.biosystems.2013.12.002 PMID:24325852

Anighoro, A., Bajorath, J., & Rastelli, G. (2014). Polypharmacology: Challenges and Opportunities in Drug Discovery. *Journal of Medicinal Chemistry*, *57*(19), 7874–7887. doi:10.1021/jm5006463 PMID:24946140

Aoyama, T., Suzuki, Y., & Ichikawa, H. (1993). Neural networks applied to structure-activity relationships. *Journal of Medicinal Chemistry*, *33*(3), 905–908. doi:10.1021/jm00165a004 PMID:2308139

Aronov, A. M. (2005). Predictive in silicomodeling for hERG channel blockers. *Drug Discovery Today*, *10*(2), 149–155. doi:10.1016/S1359-6446(04)03278-7 PMID:15718164

Baell, J. B., & Holloway, G. A. (2010). New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of Medicinal Chemistry*, *53*(7), 2719–2740. doi:10.1021/jm901137j PMID:20131845

Baroni, M., Cruciani, G., Sciabola, S., Perruccio, F., & Mason, J. S. (2007). A Common Reference Framework for Analyzing/Comparing Proteins and Ligands. Fingerprints for Ligands And Proteins (FLAP): Theory and Application. *Journal of Chemical Information and Modeling*, *47*(2), 279–294. doi:10.1021/ci600253e PMID:17381166

Barreca, M. L., De Luca, L., Iraci, N., Rao, A., Ferro, S., Maga, G., & Chimirri, A. (2007). Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors. *Journal of Chemical Information and Modeling*, *47*(2), 557–562. doi:10.1021/ci600320q PMID:17274611

Barton, M., Filardo, E. J., Lolait, S. J., Thomas, P., Maggiolini, M., & Prossnitz, E. R. (2017). Twenty years of the G protein-coupled estrogen receptor GPER: Historical and personal perspectives. *The Journal of Steroid Biochemistry and Molecular Biology*. doi:10.1016/j.jsbmb.2017.03.021 PMID:28347854

Batista, J., Hawkins, P. C., Tolbert, R., & Geballe, M. T. (2014). SiteHopper - a unique tool for binding site comparison. *Journal of Cheminformatics*, *6*(Suppl. 1), 57. doi:10.1186/1758-2946-6-S1-P57

Bauer, M. R., Ibrahim, T. M., Vogel, S. M., & Boeckler, F. M. (2013). Evaluation and Optimization of Virtual Screening Workflows with DEKOIS 2.0 – A Public Library of Challenging Docking Benchmark Sets. *Journal of Chemical Information and Modeling*, *53*(6), 1447–1462. doi:10.1021/ci400115b PMID:23705874

Beisken, S., Meinl, T., Wiswedel, B., de Figueiredo, L. F., Berthold, M., & Steinbeck, C. (2013). KNIME-CDK: Workflow-driven cheminformatics. *BMC Bioinformatics*, *14*(1), 257. doi:10.1186/1471-2105-14-257 PMID:24103053

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Berenger, F., Voet, A., Lee, X. Y., & Zhang, K. Y. (2014). A rotation-translation invariant molecular descriptor of partial charges and its use in ligand-based virtual screening. *Journal of Cheminformatics*, *6*(23). doi:10.1186/1758-2946-6-23 PMID:24887178

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... Bourne, E. P. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. doi:10.1093/nar/28.1.235 PMID:10592235

Bleeker, S. E., Moll, H. A., Steyerberg, E. W., Donders, A. R. T., Derksen-Lubsen, G., Grobbee, D. E., & Moons, K. G. M. (2003). External validation is necessary in prediction research: A clinical example. *Journal of Clinical Epidemiology*, *56*(9), 826–832. doi:10.1016/S0895-4356(03)00207-5 PMID:14505766

Blum, L. C., & Reymond, J.-L. (2009). 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *Journal of the American Chemical Society*, *131*(25), 8732–8733. doi:10.1021/ja902302h PMID:19505099

Böhm, H. J. (1994). The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of Computer-Aided Molecular Design*, *8*(3), 243–256. doi:10.1007/BF00126743 PMID:7964925

Bordogna, A., Pandini, A., & Bonati, L. (2011). Predicting the Accuracy of Protein–Ligand Docking on Homology Models. *Journal of Computational Chemistry*, *32*(1), 81–98. doi:10.1002/jcc.21601 PMID:20607693

Bringmann, G., & Rummey, C. (2003). 3D QSAR investigations on antimalarial naphthylisoquinoline alkaloids by comparative molecular similarity indices analysis (CoMSIA), based on different alignment approaches. *Journal of Chemical Information and Computer Sciences*, *43*(1), 304–316. doi:10.1021/ci025570s PMID:12546566

Castillo-González, D., Mergny, J. L., De Rache, A., Pérez-Machado, G., Cabrera-Pérez, M. A., Nicolotti, O., ... Cruz-Monteagudo, M. (2015). Harmonization of QSAR best practices and molecular docking provides an efficient virtual screening tool for discovering new G-quadruplex ligands. *Journal of Chemical Information and Modeling*, *55*(10), 2094–2110. doi:10.1021/acs.jcim.5b00415 PMID:26355653

Cavasotto, C. N., & Phatak, S. S. (2009). Homology modeling in drug discovery: Current trends and applications. *Drug Discovery Today*, *14*(13–14), 676–683. doi:10.1016/j.drudis.2009.04.006 PMID:19422931

Chang, C., Jeyachandran, S., Hu, N. J., Liu, C.-L., Lin, S.-Y., Wang, Y.-S., ... Hou, M. H. (2015). Structure-based virtual screening and experimental validation of the discovery of inhibitors targeted towards the human coronavirus nucleocapsid protein. *Molecular BioSystems*, *12*(1), 59–66. doi:10.1039/C5MB00582E PMID:26542199

Chartier, M., Adriansen, E., & Najmanovich, R. (2016). IsoMIF Finder: Online detection of binding site molecular interaction field similarities. *Bioinformatics (Oxford, England)*, *32*(4), 621–623. doi:10.1093/bioinformatics/btv616 PMID:26504139

Chaudhari, R., Tan, Z., Huang, B., & Zhang, S. (2017). Computational polypharmacology: A new paradigm for drug discovery. *Expert Opinion on Drug Discovery*, *12*(3), 279–291. doi:10.1080/17460441.2017.1280024 PMID:28067061

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chen, J., Swamidass, S. J., Dou, Y., Bruand, J., & Baldi, P. (2005). ChemDB: A public database of small molecules and related chemoinformatics resources. *Bioinformatics (Oxford, England)*, *21*(22), 4133–4139. doi:10.1093/bioinformatics/bti683 PMID:16174682

Chen, M., Vijay, V., Shi, Q., Liu, Z., Fang, H., & Tong, W. (2011). FDA-approved drug labeling for the study of drug-induced liver injury. *Drug Discovery Today*, *16*(15–16), 697–703. doi:10.1016/j.drudis.2011.05.007 PMID:21624500

Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., ... Tropsha, A. (2014). QSAR modeling: Where have you been? Where are you going to? *Journal of Medicinal Chemistry*, *57*(12), 4977–5010. doi:10.1021/jm4004285 PMID:24351051

Cohen, P. G. (2008). Obesity in men: The hypogonadal–estrogen receptor relationship and its effect on glucose homeostasis. *Medical Hypotheses*, *70*(2), 358–360. doi:10.1016/j.mehy.2007.05.020 PMID:17825496

Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, *110*(18), 5959–5967. doi:10.1021/ja00226a005 PMID:22148765

Crunkhorn, S. (2017). Computational chemistry: Novel virtual screening approach. *Nature Reviews. Drug Discovery*, *16*(1), 18–18. doi:10.1038/nrd.2016.272 PMID:28031577

Cumming, J. G., Davis, A. M., Muresan, S., Haeberlein, M., & Chen, H. (2013). Chemical predictive modelling to improve compound quality. *Nature Reviews. Drug Discovery*, *12*(12), 948–962. doi:10.1038/nrd4128 PMID:24287782

Dahlin, J. L., Nissink, J. W. M., Strasser, J. M., Francis, S., Higgins, L., Zhou, H., ... Walters, M. A. (2015). PAINS in the Assay: Chemical Mechanisms of Assay Interference and Promiscuous Enzymatic Inhibition Observed during a Sulfhydryl-Scavenging HTS. *Journal of Medicinal Chemistry*, *58*(5), 2091–2113. doi:10.1021/jm5019093 PMID:25634295

Del Rio, A. D., Barbosa, A. J. M., Caporuscio, F., & Mangiatordi, G. F. (2010). CoCoCo: A free suite of multiconformational chemical databases for high-throughput virtual screening purposes. *Molecular BioSystems*, *6*(11), 2122–2128. doi:10.1039/c0mb00039f PMID:20694263

Devillers, J. (2004). Linear versus nonlinear QSAR modeling of the toxicity of phenol derivatives to Tetrahymenapyriformis. *SAR and QSAR in Environmental Research*, *15*(4), 237–249. doi:10.1080/10629360410001724905 PMID:15370415

DeWitte, R. S., & Shakhnovich, E. I. (1996). SMoG: De Novo Design Method Based on Simple, Fast, and Accurate Free Energy Estimates. 1. Methodology and Supporting Evidence. *Journal of the American Chemical Society*, *118*(47), 11733–11744. doi:10.1021/ja960751u

Ding, B., Wang, J., Li, N., & Wang, W. (2013). Characterization of small molecule binding. I. Accurate identification of strong inhibitors in virtual screening. *Journal of Chemical Information and Modeling*, *53*(1), 114–122. doi:10.1021/ci300508m PMID:23259763

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Donati, M., Conforti, A., Lenti, M. C., Capuano, A., Bortolami, O., Motola, D., ... Leone, R. (2016). Risk of acute and serious liver injury associated to nimesulide and other NSAIDs: Data from drug-induced liver injury case-control study in Italy. *British Journal of Clinical Pharmacology*, *82*(1), 238–248. doi:10.1111/bcp.12938 PMID:26991794

Douali, L., Villemin, D., & Cherqaoui, D. (2003). Neural Networks: Accurate Nonlinear QSAR Model for HEPT Derivatives. *Journal of Chemical Information and Computer Sciences*, *43*(4), 1200–1207. doi:10.1021/ci034047q PMID:12870912

Dror, O., Shulman-Peleg, A., Nussinov, R., & Wolfson, H. J. (2004). Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. *Current Medicinal Chemistry*, *11*(1), 71–90. doi:10.2174/0929867043456287 PMID:14754427

Egan, W. J., Merz, K. M., & Baldwin, J. J. (2000). Prediction of drug absorption using multivariate statistics. *Journal of Medicinal Chemistry*, *43*(21), 3867–3877. doi:10.1021/jm000292e PMID:11052792

Everts, M., Cihlar, T., Bostwick, J. R., & Whitley, R. J. (2017). Accelerating Drug Development: Antiviral Therapies for Emerging Viruses as a Model. *Annual Review of Pharmacology and Toxicology*, *57*(1), 155–169. doi:10.1146/annurev-pharmtox-010716-104533 PMID:27483339

Ferrari, A. M., Wei, B. Q., Costantino, L., & Shoichet, B. K. (2004). Soft Docking and Multiple Receptor Conformations in Virtual Screening. *Journal of Medicinal Chemistry*, *47*(21), 5076–5084. doi:10.1021/jm049756p PMID:15456251

Flick, J., Tristram, F., & Wenzel, W. (2012). Modeling loop backbone flexibility in receptor-ligand docking simulations. *Journal of Computational Chemistry*, *33*(31), 2504–2515. doi:10.1002/jcc.23087 PMID:22886372

Floris, M., Manganaro, A., Nicolotti, O., Medda, R., Mangiatordi, G. F., & Benfenati, E. (2014). A generalizable definition of chemical similarity for read-across. *Journal of Cheminformatics*, *6*(1), 39. doi:10.118613321-014-0039-1 PMID:25383097

Fontana, R. J., Watkins, P. B., Bonkovsky, H. L., Chalasani, N., Davern, T., Serrano, J., & Rochon, J.DILIN Study Group. (2009). Drug-Induced Liver Injury Network (DILIN) prospective study: Rationale, design and conduct. *Drug Safety*, *32*(1), 55–68. doi:10.2165/00002018-200932010-00005 PMID:19132805

Free, S. M., & Wilson, J. W. (1964). A mathematical contribution to structure-activity studies. *Journal of Medicinal Chemistry*, *7*(4), 395–399. doi:10.1021/jm00334a001 PMID:14221113

Freire, E. (2008). Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today*, *13*(19–20), 869–874. doi:10.1016/j.drudis.2008.07.005 PMID:18703160

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., ... Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, *47*(7), 1739–1749. doi:10.1021/jm0306430 PMID:15027865

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Frimurer, T. M., Bywater, R., Nærum, L., Lauritsen, L. N., & Brunak, S. (2000). Improving the Odds in Discriminating Drug-like from Non Drug-like Compounds. *Journal of Chemical Information and Computer Sciences*, *40*(6), 1315–1324. doi:10.1021/ci0003810 PMID:11128089

Fujita, T. (1964). C. p- σ- π Analysis.A method for the correlation of biological activity and chemical structure. *Journal of the American Chemical Society*, *86*(8), 1616–1626. doi:10.1021/ja01062a035

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., ... Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, *40*(Database issue), D1100–D1107. doi:10.1093/nar/gkr777 PMID:21948594

Geladi, P., & Kowalski, B. R. (1986). Partial Least-Squares Regression: A Tutorial. *Analytica Chimica Acta*, *185*, 1–17. doi:10.1016/0003-2670(86)80028-9

Ghaffar, A., Batool, S., Mushtaq, G., & Kamal, M. A. (2017). In silico Analysis of AMP-Activated Protein Kinase and Ligand-based Virtual Screening for Identification of Novel AMPK Activators. *Current Computer-aided Drug Design*, *13*(3). doi:10.2174/1573409913666170309144722 PMID:28294049

Gillet, V. J., & Nicolotti, O. (2000). Evaluation of reactant-based and product-based approaches to the design of combinatorial libraries. *Perspectives in Drug Discovery and Design*, *20*(1), 265–287. doi:10.1023/A:1008797526431

Glaab, E. (2016). Building a virtual ligand screening pipeline using free software: A survey. *Briefings in Bioinformatics*, *17*(2), 352–366. doi:10.1093/bib/bbv037 PMID:26094053

Goecks, J., Nekrutenko, A., Taylor, J., & Galaxy Team, T. (2010). Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*, *11*(8), R86. doi:10.1186/gb-2010-11-8-r86 PMID:20738864

Gohlke, H., Hendlich, M., & Klebe, G. (2000). Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, *295*(2), 337–356. doi:10.1006/jmbi.1999.3371 PMID:10623530

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning* (1st ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

González, P. M., Acharya, C., Mackerell, A. D., & Polli, J. E. (2009). Inhibition requirements of the human apical sodium-dependent bile acid transporter (hASBT) using aminopiperidine conjugates of glutamyl-bile acids. *Pharmaceutical Research*, *26*(7), 1665–1678. doi:10.100711095-009-9877-3 PMID:19384469

Gozalbes, R., & Pineda-Lucena, A. (2010). QSAR-based solubility model for drug-like compounds. *Bioorganic & Medicinal Chemistry*, *18*(19), 7078–7084. doi:10.1016/j.bmc.2010.08.003 PMID:20810286

Gruvberger, S., Ringnér, M., Chen, Y., Panavally, S., Saal, L. H., Borg, A., ... Meltzer, P. S. (2001). Estrogen Receptor Status in Breast Cancer Is Associated with Remarkably Distinct Gene Expression Patterns. *Cancer Research*, *61*(16), 5979–5984. PMID:11507038

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Guillaume, M., Handgraaf, S., Fabre, A., Raymond-Letron, I., Riant, E., Montagner, A., ... Gourdy, P. (2017). Selective Activation of Estrogen Receptor α Activation Function-1 Is Sufficient to Prevent Obesity, Steatosis, and Insulin Resistance in Mouse. *American Journal of Pathology*, *187*(6), 1273–1287. doi:10.1016/j.ajpath.2017.02.013 PMID:28502695

Hawkins, P. C. D., Skillman, A. G., Warren, G. L., Ellingson, B. A., & Stahl, M. T. (2010). Conformer generation with OMEGA: Algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *Journal of Chemical Information and Modeling*, *50*(4), 572–584. doi:10.1021/ci100031x PMID:20235588

Hillisch, A., Pineda, L. F., & Hilgenfeld, R. (2004). Utility of homology models in the drug discovery process. *Drug Discovery Today*, *9*(15), 659–669. doi:10.1016/S1359-6446(04)03196-4 PMID:15279849

Hou, X., Li, K., Yu, X., Sun, J., & Fang, H. (2015). Protein Flexibility in Docking-Based Virtual Screening: Discovery of Novel Lymphoid-Specific Tyrosine Phosphatase Inhibitors Using Multiple Crystal Structures. *Journal of Chemical Information and Modeling*, *55*(9), 1973–1983. doi:10.1021/acs.jcim.5b00344 PMID:26360643

Houston, D. R., & Walkinshaw, M. D. (2013). Consensus Docking: Improving the Reliability of Docking in a Virtual Screening Context. *Journal of Chemical Information and Modeling*, *53*(2), 384–390. doi:10.1021/ci300399w PMID:23351099

Huang, N., Shoichet, B. K., & Irwin, J. J. (2006). Benchmarking Sets for Molecular Docking. *Journal of Medicinal Chemistry*, *49*(23), 6789–6801. doi:10.1021/jm0608356 PMID:17154509

Huey, R., Morris, G. M., Olson, A. J., & Goodsell, D. S. (2007). Asemiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*, *28*(6), 1145–1152. doi:10.1002/jcc.20634 PMID:17274016

Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, *162*(6), 1239–1249. doi:10.1111/j.1476-5381.2010.01127.x PMID:21091654

Ihaka, R., & Gentleman, R. R. (1996). A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, *5*(3), 299–314.

Ihlenfeldt, W.-D., Voigt, J. H., Bienfait, B., Oellien, F., & Nicklaus, M. C. (2002). Enhanced CACTVS browser of the Open NCI Database. *Journal of Chemical Information and Computer Sciences*, *42*(1), 46–57. doi:10.1021/ci010056s PMID:11855965

InterBioScreen. (2012). Natural Compound (NC) database. Retrieved from http://www.ibscreen.com/natural.shtm

Irwin, J. J., & Shoichet, B. K. (2005). ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *Journal of Chemical Information and Modeling*, *45*(1), 177–182. doi:10.1021/ci049714+ PMID:15667143

Jalencas, X., & Mestres, J. (2013). Identification of Similar Binding Sites to Detect Distant Polypharmacology. *Molecular Informatics*, *32*(11–12), 976–990. doi:10.1002/minf.201300082 PMID:27481143

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jambon, M., Imberty, A., Deléage, G., & Geourjon, C. (2003). A new bioinformatic approach to detect common 3D sites in protein structures. *Proteins*, *52*(2), 137–145. doi:10.1002/prot.10339 PMID:12833538

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, *267*(3), 727–748. doi:10.1006/jmbi.1996.0897 PMID:9126849

Kar, S., & Roy, K. (2013). How far can virtual screening take us in drug discovery? *Expert Opinion on Drug Discovery*, *8*(3), 245–261. doi:10.1517/17460441.2013.761204 PMID:23330660

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., ... Bryant, S. H. (2016). PubChem Substance and Compound databases. *Nucleic Acids Research*, *44*(Database issue), D1202–D1213. doi:10.1093/nar/gkv951 PMID:26400175

King, R. D., Hirst, J. D., &Sternberg, M. J.E. (1993). New approaches to QSAR: Neural networks and machine learning. *Perspectives in Drug Discovery and Design, 1*, 279-290. doi:10.1007/BF02174529

Kirchmair, J., Laggner, C., Wolber, G., & Langer, T. (2005). Comparative analysis of protein-bound ligand conformations with respect to catalysts conformational space subsampling algorithms. *Journal of Chemical Information and Modeling*, *45*(2), 422–430. doi:10.1021/ci049753l PMID:15807508

Kiss, R., Kiss, B., Könczöl, Á., Szalai, F., Jelinek, I., László, V., ... Kerasu, G. M. (2008). Discovery of Novel Human Histamine H4 Receptor Ligands by Large-Scale Structure-Based Virtual Screening. *Journal of Medicinal Chemistry*, *51*(11), 3145–3153. doi:10.1021/jm7014777 PMID:18459760

Kitamura, K., Tamura, Y., Ueki, T., Ogata, K., Noda, S., Himeno, R., & Chuman, H. (2014). Binding Free-Energy Calculation Is a Powerful Tool for Drug Optimization: Calculation and Measurement of Binding Free Energy for 7-Azaindole Derivatives to Glycogen Synthase Kinase-3β. *Journal of Chemical Information and Modeling*, *54*(6), 1653–1660. doi:10.1021/ci400719v PMID:24884406

Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews. Drug Discovery*, *3*(11), 935–949. doi:10.1038/nrd1549 PMID:15520816

Klabunde, T. (2007). Chemogenomic approaches to drug discovery: Similar receptors bind similar ligands. *British Journal of Pharmacology*, *152*(1), 5–7. doi:10.1038j.bjp.0707308 PMID:17533415

Klebe, G. (2006). Virtual ligand screening: Strategies, perspectives and limitations. *Drug Discovery Today*, *11*(13–14), 580–594. doi:10.1016/j.drudis.2006.05.012 PMID:16793526

Klebe, G., Abraham, U., & Mietzner, T. (1994). Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *Journal of Medicinal Chemistry*, *37*(24), 4130–4146. doi:10.1021/jm00050a010 PMID:7990113

Koeppen, H., Kriegl, J., Lessel, U., Tautermann, C. S., & Wellenzohn, B. (2011). Ligand-Based Virtual Screening. In C. Sotriffer (Ed.), Virtual Screening (pp. 61–85). Wiley-VCH Verlag GmbH & Co. KGaA. doi:10.1002/9783527633326.ch3

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kohavi, R. (1995). A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Vol. 2, pp. 1137–1143). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Kuo, C.-L., Assefa, H., Kamath, S., Brzozowski, Z., Slawinski, J., Saczewski, F., ... Neamati, N. (2004). Application of CoMFA and CoMSIA 3D-QSAR and docking studies in optimization of mercaptobenzenesulfonamides as HIV-1 integrase inhibitors. *Journal of Medicinal Chemistry*, *47*(2), 385–399. doi:10.1021/jm030378i PMID:14711310

La, D., Esquivel-Rodríguez, J., Venkatraman, V., Li, B., Sael, L., Ueng, S., ... Kihara, D. I. (2009). 3D-SURFER: Software for high-throughput protein surface comparison and analysis. *Bioinformatics (Oxford, England)*, *25*(21), 2843–2844. Epub2009Sep16 doi:10.1093/bioinformatics/btp542 PMID:19759195

Lagorce, D., Sperandio, O., Galons, H., Miteva, M. A., & Villoutreix, B. O. (2008). FAF-Drugs2: Free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics*, *9*(1), 396. doi:10.1186/1471-2105-9-396 PMID:18816385

Lavecchia, A., & Cerchia, C. (2016). In silico methods to address polypharmacology: Current status, applications and future perspectives. *Drug Discovery Today*, *21*(2), 288–298. doi:10.1016/j.drudis.2015.12.007 PMID:26743596

Lavecchia, A., & Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: A critical review. *Current Medicinal Chemistry*, *20*(23), 2839–2860. doi:10.2174/09298673113209990001 PMID:23651302

Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., ... Wishart, D. S. (2014). DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Research*, *42*(Database issue), D1091–D1097. doi:10.1093/nar/gkt1068 PMID:24203711

Lehtovuori, P. T., & Nyrönen, T. H. (2006). SOMA – Workflow for small molecule property calculations on a multiplatform computing grid. *Journal of Chemical Information and Modeling*, *46*(2), 620–625. doi:10.1021/ci050388n PMID:16562991

Lešnik, S., Štular, T., Brus, B., Knez, D., Gobec, S., Janežič, D., & Konc, J. (2015). LiSiCA: A software for Ligand-Based Virtual Screening and its application for the discovery of Butyrylcholinesterase inhibitors. *Journal of Chemical Information and Modeling*, *55*(8), 1521–1528. doi:10.1021/acs.jcim.5b00136 PMID:26158767

Lionta, E., Spyrou, G., Vassilatis, D. K., & Cournia, Z. (2014). Structure-based virtual screening for drug discovery: Principles, applications and recent advances. *Current Topics in Medicinal Chemistry*, *14*(16), 1923–1938. doi:10.2174/1568026614666140929124445 PMID:25262799

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *46*(1–3), 3–26. doi:10.1016/S0169-409X(00)00129-0 PMID:11259830

Lu, Q., Hao, P., Curcin, V., He, W., Li, Y.-Y., Luo, Q.-M., ... Li, Y.-X. (2006). KDE Bioscience: Platform for bioinformatics analysis workflows. *Journal of Biomedical Informatics*, *39*(4), 440–450. doi:10.1016/j.jbi.2005.09.001 PMID:16260186

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lyne, P. D. (2002). Structure-based virtual screening: An overview. *Drug Discovery Today*, *7*(20), 1047–1055. doi:10.1016/S1359-6446(02)02483-2 PMID:12546894

Malmstrom, R. D., & Watowich, S. J. (2011). Using Free Energy of Binding Calculations To Improve the Accuracy of Virtual Screening Predictions. *Journal of Chemical Information and Modeling*, *51*(7), 1648–1655. doi:10.1021/ci200126v PMID:21696204

Mangiatordi, G. F., Alberga, D., Altomare, C. D., Carotti, A., Catto, M., Cellamare, S., ... Nicolotti, O. (2016). Mind the Gap! A Journey towards Computational Toxicology. *Molecular Informatics*, *35*(8–9), 294–308. doi:10.1002/minf.201501017 PMID:27546034

Mangiatordi, G.F., Alberga, D., Siragusa, L., Goracci, L., Lattanzi, G., & Nicolotti, O. (2015). Challenging AQP4 druggabilityfor NMO-IgG antibody binding using Molecular Dynamics and Molecular Interaction Fields. *Biochimica et BiophysicaActa - Biomembranes, 1848*(7), 1462–1471. doi:. doi:10.1016/j.bbamem.2015.03.019

Mangiatordi, G. F., Alberga, D., Trisciuzzi, D., Lattanzi, G., & Nicolotti, O. (2016). Human aquaporin-4 and molecular modeling: Historical perspective and view to the future. *International Journal of Molecular Sciences*, *17*(7), 1119. doi:10.3390/ijms17071119 PMID:27420052

Mangiatordi, G. F., Trisciuzzi, D., Alberga, D., Denora, N., Iacobazzi, R. M., Gadaleta, D., ... Nicolotti, O. (2017). Novel chemotypes targeting tubulin at the Colchicine Binding Site and unbiasing P-glycoprotein. *European Journal of Medicinal Chemistry*. doi:10.1016/j.ejmech.2017.07.037

Mansouri, K., Abdelaziz, A., Rybacka, A., Roncaglioni, A., Tropsha, A., Varnek, A., ... Judson, R. S. (2016). CERAPP: Collaborative Estrogen Receptor Activity Prediction Project. *Environmental Health Perspectives*, *124*(7), 1023–1033. doi:10.1289/ehp.1510267 PMID:26908244

Martinaityte, I., Jorde, R., Emaus, N., Eggen, A. E., Joakimsen, R. M., & Kamycheva, E. (2017). Bone mineral density is associated with vitamin D related rs6013897 and estrogen receptor polymorphism rs4870044: The Tromsø study. *PLoS ONE*, *12*(3), e0173045. doi:10.1371/journal.pone.0173045 PMID:28253304

Mazanetz, M. P., Marmon, R. J., Reisser, C. B. T., & Morao, I. (2012). Drug discovery applications for KNIME: An open source data mining platform. *Current Topics in Medicinal Chemistry*, *12*(18), 1965–1979. doi:10.2174/156802612804910331 PMID:23110532

McInnes, C. (2007). Virtual screening strategies in drug discovery. *Current Opinion in Chemical Biology*, *11*(5), 494–502. doi:10.1016/j.cbpa.2007.08.033 PMID:17936059

Melville, J. L., Burke, E. K., & Hirst, J. D. (2009). Machine learning in virtual screening. *Combinatorial Chemistry & High Throughput Screening*, *12*(4), 332–343. doi:10.2174/138620709788167980 PMID:19442063

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092. doi:10.1063/1.1699114

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mishra, V., & Siva Prasad, C. V. S. (2011). Ligand based virtual screening to find novel inhibitors against plant toxin Ricin by using the ZINC database. *Bioinformation*, *7*(2), 46–51. doi:10.6026/97320630007046 PMID:21938204

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry, 19*(14), 1639–1662. doi:<1639::AID JCC10>3.0.CO;2-B doi:10.1002/(SICI)1096-987X(19981115)19:14

Myers, S., & Baker, A. (2001). Drug discoveryan operating model for a new era. *Nature Biotechnology*, *19*(8), 727–730. doi:10.1038/90765 PMID:11479559

Mysinger, M. M., Carchia, M., Irwin, J. J., & Shoichet, B. K. (2012). Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *Journal of Medicinal Chemistry*, *55*(14), 6582–6594. doi:10.1021/jm300687e PMID:22716043

Najmanovich, R. J. (2017). Evolutionary studies of ligand binding sites in proteins. *Current Opinion in Structural Biology*, *45*, 85–90. doi:10.1016/j.sbi.2016.11.024 PMID:27992825

Nicolotti, O., Benfenati, E., Carotti, A., Gadaleta, D., Gissi, A., Mangiatordi, G. F., & Novellino, E. (2014). REACH and in silico methods: An attractive opportunity for medicinal chemists. *Drug Discovery Today*, *19*(11), 1757–1768. doi:10.1016/j.drudis.2014.06.027 PMID:24998783

Nicolotti, O., & Carotti, A. (2006). QSAR and QSPR studies of a highly structured physicochemical domain. *Journal of Chemical Information and Modeling*, *46*(1), 264–276. doi:10.1021/ci0502931 PMID:16426062

Nicolotti, O., Giangreco, I., Miscioscia, T. F., & Carotti, A. (2009). Improving quantitative structure-activity relationships through multiobjective optimization. *Journal of Chemical Information and Modeling*, *49*(10), 2290–2302. doi:10.1021/ci9002409 PMID:19785453

Nicolotti, O., Miscioscia, T.F., Carotti, A., Leonetti, F., & Carotti, A. (2008). An integrated approach to ligand- and structure-based drug design: Development and application to a series of serine protease inhibitors. *Journal of Chemical Information and Modeling, 48*(6), 1211-1226, . doi:10.1021/ci800015s

Noureen, N., Kalsoom, S., & Rashid, H. (2010). Ligand based pharmacophore modelling of anticancer histone deacetylase inhibitors. *African Journal of Biotechnology*, *9*(25), 3923–3931.

Odolczyk, N., Fritsch, J., Norez, C., Servel, N., Cunha, M. F., da, & Bitam, S. (2013). Discovery of novel potent ΔF508-CFTR correctors that target the nucleotide binding domain. *EMBO Molecular Medicine*, *5*(10), 1484–1501. doi:10.1002/emmm.201302699 PMID:23982976

Oinn, T., Addis, M., Ferris, J., Marvin, D., Senger, M., Greenwood, M., ... Li, P. (2004). Taverna: A tool for the composition and enactment of bioinformatics workflows. *Bioinformatics (Oxford, England)*, *20*(17), 3045–3054. doi:10.1093/bioinformatics/bth361 PMID:15201187

Onodera, K., Satou, K., & Hirota, H. (2007). Evaluations of Molecular Docking Programs for Virtual Screening. *Journal of Chemical Information and Modeling*, *47*(4), 1609–1618. doi:10.1021/ci7000378 PMID:17602548

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Oprea, T. (2000). Property distribution of drug-related chemical databases. *Journal of Computer-Aided Molecular Design*, *14*(3), 251–264. doi:10.1023/A:1008130001697 PMID:10756480

Oprea, T. I. (2002). Virtual Screening in Lead Discovery: A Viewpoint. *Molecules (Basel, Switzerland)*, *7*(1), 51–62. doi:10.3390/70100051

Oshiro, C., Bradley, E. K., Eksterowicz, J., Evensen, E., Lamb, M. L., Lanctot, J. K., ... Grootenhuis, P. D. J. (2004). Performance of 3D-database molecular docking studies into homology models. *Journal of Medicinal Chemistry*, *47*(3), 764–767. doi:10.1021/jm0300781 PMID:14736258

Pieper, U., Webb, B. M., Dong, G. Q., Schneidman-Duhovny, D., Fan, H., Kim, S. J., ... Sali, A. (2014). ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, *42*, D336–D346. doi:10.1093/nar/gkt1144 PMID:24271400

Pisani, F., Mola, M. G., Simone, L., Rosito, S., Alberga, D., Mangiatordi, G. F., ... Nicchia, G. P. (2014). Identification of a point mutation impairing the binding between aquaporin-4 and neuromyelitisoptica autoantibodies. *The Journal of Biological Chemistry*, *289*(44), 30578–30589. doi:10.1074/jbc.M114.582221 PMID:25239624

Qing, X., Lee, X. Y., Raeymaecker, J. D., Tame, J. R., Zhang, K. Y., Maeyer, M. D., & Voet, A. R. (2014, November 11). Pharmacophore modeling: advances, limitations, and current utility in drug discovery. *Journal of Receptor, Ligand and Channel Research*. doi:10.2147/JRLCR.S46843

Rainsford, K. D. (2006). Nimesulide - a multifactorial approach to inflammation and pain: Scientific and clinical consensus. *Current Medical Research and Opinion*, *22*(6), 1161–1170. doi:10.1185/030079906X104849 PMID:16846549

Reddy, A. S., Pati, S. P., Kumar, P. P., Pradeep, H. N., & Sastry, G. N. (2007). Virtual Screening in Drug Discovery - A Computational Perspective. *Current Protein & Peptide Science*, *8*(4), 329–351. doi:10.2174/138920307781369427 PMID:17696867

Ren, J., Xie, L., Li, W. W., & Bourne, P. E. (2010). SMAP-WS: a parallel web service for structural proteome-wide ligand-binding site comparison. *Nucleic Acids Research, 38*(Suppl. 2), W441–W444. doi:10.1093/nar/gkq400

Ripphausen, P., Nisius, B., & Bajorath, J. (2011). State-of-the-art in ligand-based virtual screening. *Drug Discovery Today*, *16*(9–10), 372–376. doi:10.1016/j.drudis.2011.02.011 PMID:21349346

Rogers, D., & Hopfinger, A. J. (1993). Application of Genetic Function Approximation to Quantitative Structure-Activity Relationships and Quantitative Structure-Property Relationships. *Journal of Chemical Information and Computer Sciences*, *34*(4), 854–866. doi:10.1021/ci00020a020

Rognan, D. (2017). The impact of in silico screening in the discovery of novel and safer drug candidates. *Pharmacology & Therapeutics*, *175*, 47–66. doi:10.1016/j.pharmthera.2017.02.034 PMID:28223231

Roy, K., Kar, S., & Das, R. N. (2015). *Statistical Methods in QSAR/QSPR, A Primer on QSAR/QSPR Modeling*. Springer Briefs in Molecular Science. doi:10.1007/978-3-319-17281-1_2

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ruddigkeit, L., van Deursen, R., Blum, L. C., & Reymond, J.-L. (2012). Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. *Journal of Chemical Information and Modeling*, *52*(11), 2864–2875. doi:10.1021/ci300415d PMID:23088335

Ruiz-Carmona, S., Alvarez-Garcia, D., Foloppe, N., Garmendia-Doval, A. B., Juhos, S., Schmidtke, P., ... Morley, S. D. (2014). rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Computational Biology*, *10*(4), e1003571. doi:10.1371/journal.pcbi.1003571 PMID:24722481

Schellhammer, I., & Rarey, M. (2004). FlexX-Scan: Fast, structure-based virtual screening. *Proteins*, *57*(3), 504–517. doi:10.1002/prot.20217 PMID:15382244

Schrodinger, LLC. (2012). QikProp Version 3.2. Retrieved from http://www.schrodinger.com

Seifert, M. H. J. (2009). Targeted scoring functions for virtual screening. *Drug Discovery Today*, *14*(11–12), 562–569. doi:10.1016/j.drudis.2009.03.013 PMID:19508918

Seifert, M. H. J., & Lang, M. (2008). Essential Factors for Successful Virtual Screening. *Mini Reviews in Medicinal Chemistry*, *8*(1), 63–72. doi:10.2174/138955708783331540 PMID:18220986

Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature*, *432*(7019), 862–865. doi:10.1038/nature03197 PMID:15602552

Siragusa, L., Cross, S., Baroni, M., Goracci, L., & Cruciani, G. (2015). BioGPS: Navigating biological space to predict polypharmacology, off-targeting, and selectivity. *Proteins: Structure, Function, and Bioinformatics*, *83*(3), 517–532. doi:10.1002/prot.24753 PMID:25556939

Sirois, S., Hatzakis, G., Wei, D., Du, Q., & Chou, K.-C. (2005). Assessment of chemical libraries for their druggability. *Computational Biology and Chemistry*, *29*(1), 55–67. doi:10.1016/j.compbiolchem.2004.11.003 PMID:15680586

Sleeman, D., & Edwards, P. (Eds.), (1992). *Proceedings of the Ninth International Workshop (ML92)*. Morgan Kaufmann.

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). Computational Methods in Drug Discovery. *Pharmacological Reviews*, *66*(1), 334–395. doi:10.1124/pr.112.007336 PMID:24381236

Spadaro, A., Negri, M., Marchais-Oberwinkler, S., Bey, E., & Frotscher, M. (2012). Hydroxybenzothiazoles as New Nonsteroidal Inhibitors of 17β-Hydroxysteroid Dehydrogenase Type 1 (17β-HSD1). *PLoS ONE*, *7*(1), e29252. doi:10.1371/journal.pone.0029252 PMID:22242164

Sterling, T., & Irwin, J. J. (2015). ZINC 15 – Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling*, *55*(11), 2324–2337. doi:10.1021/acs.jcim.5b00559 PMID:26479676

Svetnik, V., Liaw, A., Tong, C. J., Culberson, C., Sheridan, R. P., & Feuston, B. P. (2003). Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of Chemical Information and Computer Sciences*, *43*(6), 1947–1958. doi:10.1021/ci034160g PMID:14632445

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tao, P., & Lai, L. (2001). Protein ligand docking based on empirical method for binding affinity estimation. *Journal of Computer-Aided Molecular Design*, *15*(5), 429–446. doi:10.1023/A:1011188704521 PMID:11394737

Trisciuzzi, D., Alberga, D., Mansouri, K., Judson, R., Cellamare, S., Catto, M., ... Nicolotti, O. (2015). Docking-based classification models for exploratory toxicology studies on high-quality estrogenic experimental data. *Future Medicinal Chemistry*, *7*(14), 1921–1936. doi:10.4155/fmc.15.103 PMID:26440057

Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, *45*(12), 2615–2623. doi:10.1021/jm020017n PMID:12036371

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., & Taylor, R. D. (2003). Improved protein-ligand docking using GOLD. *Proteins*, *52*(4), 609–623. doi:10.1002/prot.10465 PMID:12910460

Villoutreix, B. O., Renault, N., Lagorce, D., Sperandio, O., Montes, M., & Miteva, M. A. (2007). Free resources to assist structure-based virtual ligand screening experiments. *Current Protein & Peptide Science*, *8*(4), 381–411. doi:10.2174/138920307781369391 PMID:17696871

Wang, R., Lai, L., & Wang, S. (2002). Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design*, *16*(1), 11–26. doi:10.1023/A:1016357811882 PMID:12197663

Wang, R., Lu, Y., & Wang, S. (2003). Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *Journal of Medicinal Chemistry*, *46*(12), 2287–2303. doi:10.1021/jm0203783 PMID:12773034

Wang, R., & Wang, S. (2001). How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *Journal of Chemical Information and Computer Sciences*, *41*(5), 1422–1426. doi:10.1021/ci010025x PMID:11604043

Wang, Y., Bryant, S. H., Cheng, T., Wang, J., Gindulyte, A., Shoemaker, B. A., ... Zhang, J. (2017). PubChemBioAssay: 2017 update. *Nucleic Acids Research*, *45*(D17), D955–D963. doi:10.1093/nar/gkw1118 PMID:27899599

Weiss, S. M., & Kulikowski, C. A. (1991). *Computer systems that learn*. San Mateo, CA: Morgan Kaufmann.

Williams, A. J. (2011). Chemspider: A Platform for Crowdsourced Collaboration to Curate Data Derived From Public Compound Databases. In S. Ekins, Maggie A. Z. Hupcey, & A. J. Williams (Eds.), Collaborative Computational Technologies for Biomedical Research (pp. 363–386). John Wiley & Sons, Inc. doi:10.1002/9781118026038.ch22

Wingerchuk, D. M., Lennon, V. A., Lucchinetti, C. F., Pittock, S. J., & Weinshenker, B. G. (2007). The spectrum of neuromyelitisoptica. *Lancet Neurology*, *6*(9), 805–815. doi:10.1016/S1474-4422(07)70216-8 PMID:17706564

Wójcikowski, M., Ballester, P. J., & Siedlecki, P. (2017). Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*, *7*, 46710. doi:10.1038rep46710 PMID:28440302

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Wolber, G., & Langer, T. (2005). LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling*, *45*(1), 160–169. doi:10.1021/ci049885e PMID:15667141

Xie, L., & Bourne, P. E. (2008). Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(14), 5441–5446. doi:10.1073/pnas.0704422105 PMID:18385384

Xing, L., McDonald, J. J., Kolodziej, S. A., Kurumbail, R. G., Williams, J. M., Warren, C. J., ... Roberds, S. L. (2011). Discovery of Potent Inhibitors of Soluble Epoxide Hydrolase by Combinatorial Library Design and Structure-Based Virtual Screening. *Journal of Medicinal Chemistry*, *54*(5), 1211–1222. doi:10.1021/jm101382t PMID:21302953

Xiong, X., Yuan, H., Zhang, Y., Xu, J., Ran, T., Liu, H., ... Chen, Y. (2015). Protein flexibility oriented virtual screening strategy for JAK2 inhibitors. *Journal of Molecular Structure*, *1097*, 136–144. doi:10.1016/j.molstruc.2015.05.007

Xiong, Y., Esquivel-Rodriguez, J., Sael, L., & Kihara, D. (2014). 3D-SURFER 2.0: Web platform for real-time search and characterization of protein surfaces. *Methods in Molecular Biology (Clifton, N.J.)*, *1137*, 105–117. doi:10.1007/978-1-4939-0366-5_8 PMID:24573477

Yang, S.-Y. (2010). Pharmacophoremodeling and applications in drug discovery: Challenges and recent advances. *Drug Discovery Today*, *15*(11–12), 444–450. doi:10.1016/j.drudis.2010.03.013 PMID:20362693

Zhong, S., Zhang, Y., & Xiu, Z. (2010). Rescoring ligand docking poses. *Current Opinion in Drug Discovery & Development*, *13*(3), 326–334. PMID:20443166

Zhou, Z., Gong, Q., Epstein, M. L., & January, C. T. (1998). HERG Channel Dysfunction in Human Long QT Syndrome. Intracellular transport and functional defects. *The Journal of Biological Chemistry*, *273*(33), 21061–21066. doi:10.1074/jbc.273.33.21061 PMID:9694858

This research was previously published in the International Journal of Quantitative Structure-Property Relationships (IJQSPR), 3(1); edited by Kunal Roy; pages 134-160; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 12
Relationships between Models of Genetic Regulatory Networks with Emphasis on Discrete State Stochastic Models

Randip Pal
Texas Tech University, USA

ABSTRACT

Genetic Regulatory Networks (GRNs) represent the interconnections between genomic entities that govern the regulation of gene expression. GRNs have been represented by various types of mathematical models that capture different aspects of the biological system. This chapter discusses the relationships among the most commonly used GRN models that can enable effective integration of diverse types of sub-models. A detailed model in the form of stochastic master equation is described, followed by it coarse-scale and deterministic approximations in the form of Probabilistic Boolean Networks and Ordinary Differential Equation models respectively.

INTRODUCTION

Genetic Regulatory Networks (GRNs) represent the interconnections between genomic entities that govern the regulation of gene expression. Since biological regulatory networks are extremely detailed with numerous interactions, a single mathematical model to represent the whole biological regulatory system is generally not feasible. Depending on the purpose of modeling, the mathematical model representing the GRN brings in a level of abstraction. The focus of the modeling can be capturing interactions between RNA expressions, protein-protein interactions or interactions between metabolites. Usually, only parts of the regulome (i.e. genes, proteins and metabolites involved in gene regulation) such as transcription factors, enhancers, microRNA etc. are made explicit in a mathematical model of a GRN.

DOI: 10.4018/978-1-7998-1204-3.ch012

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

To provide further background on genetic regulatory network modeling, we present broad classifications of the models that are commonly used to capture the relationships between genomic entities. GRN models usually belong to either static or dynamical models. Static models are concerned with providing the topology of the network (such as the connectivity of various genomic entities) without attempting to model the dynamics of the interactions. Dynamic models on the other hand not only capture the connectivity of the network entities but also provide the time dependent functional relationships that govern the interactions. Commonly used pathway models as presented in databases such as KEGG and StringDB belong to the category of static models as they primarily capture the connectivity structure without the network dynamics. Examples of dynamic models include ordinary differential equation models to model the average concentration of mRNA expressions (Kampen, 1991) or Dynamic Bayesian Network (Murphy, 2002) models. Based on the model complexity involving inference of connectivity and time dependent relationships, the experimental data requirements for elucidating dynamic models are significantly higher as compared to static models.

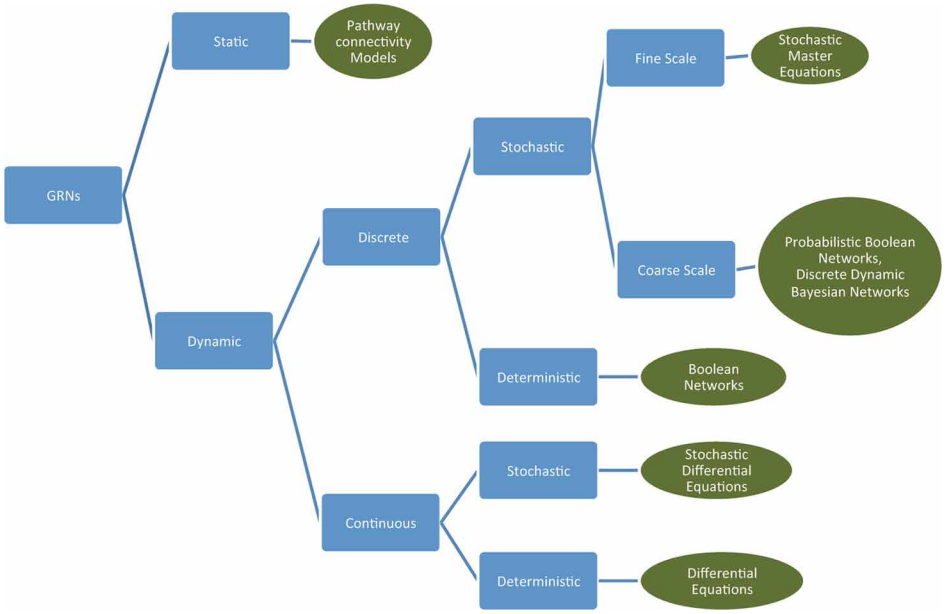
To further explore various levels of time dependent behavior of genes or proteins, dynamic network models can be essentially categorized as discrete where the amount of genomic entity is modeled as discrete values or continuous where the genomic entities are models as continuous variables. Since the dynamics of biological entities consist of discrete changes based on the number of molecules created or destroyed, a fine-scale discrete model can potentially faithfully capture the dynamics of the genomic entities. Note that we have categorized the dynamic models as discrete or continuous based on the representation of the genomic entity but a similar categorization can be applied based on discrete or continuous representation of time. Some models such as stochastic master equations (also commonly known as Chemical Master Equations) incorporate time as a continuous quantity whereas models like Boolean Networks update states at discrete time intervals. The next level of characterizations is based on whether the models considered are stochastic or deterministic. A number of studies have shown that the generation of mRNA or protein expression is stochastic in nature (Arkin, Ross, & McAdams, 1998) and thus a stochastic model is suitable to capture the true change in state probability distribution with time. However, available experimental data may limit the inference of the parameters of a stochastic model. For instance, common genomic characterizations using microarrays provide an average observation of the mRNA expressions and are not suitable for inferring the parameters of a detailed stochastic model.

Thus, deterministic models such as differential equation models are often used to represent the average behavior of the biological system (De Jong, 2001) or in other words modeling moments of the distribution rather than the detailed distribution. Dynamic models can be further classified as detailed fine-scale models or coarse-scale models where the detail refers to the level of discretization involved in modeling the gene or protein expression. For instance, our interest may be in modeling the up or down regulation of genes and thus a coarse-scale Boolean Network (Kauffman S. A., 1993; Kauffman S., 1969) with ON and OFF states for each gene will be adequate. However, if our interest is in capturing the individual mRNA molecule generation or degradation, a detailed fine-scale model such as stochastic master equations (Gillespie D. T., 1977; Pal & Bhattacharya, 2010) with an enormous state space will be required. The above categorizations are presented in a graphical form in Figure 1.

Based on the above discussion, we observe that fine-scale stochastic discrete dynamical models such as stochastic master equations can provide a detailed representation of the dynamic interactions but maybe limited by the enormous computational complexity involved in its simulation and huge data requirements for estimation of its parameters. This chapter will present a review of stochastic master equation modeling followed by some of its low complexity approximations with focus on reduced state

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Categorization of commonly used models of genetic regulatory networks (GRNs)



discrete time Probabilistic Boolean Networks. The emphasis will be on the properties maintained by the reduced order models as compared to the original model.

Chemical Master Equation Models

Chemical master equations (CMEs) are considered to be one of the detailed fundamental models for representing dynamic genomic interactions (Pal & Bhattacharya 2010). CMEs model the creation and decay of discrete molecules of involved genomic entities in a stochastic fashion. CME based models have been shown to accurately model processes in bacteriophage lambda that control the switchover from lysogenic to lytic modes of growth (Shea & Ackers, 1985) or model the dynamics of phosphorylation-dephosphorylation cycles with autophosphorylation (Bishop & Qian, 2010) .

To explain the CME models for genetic regulatory networks, let us consider a system with n molecular species and m different reaction channels where the state of the system is defined by $x = [\varphi_1, \varphi_2, ..., \varphi_n]$, $x \in N^n$ is a vector of integers representing a specific population of each of the n molecular species. For such a system, given the probability density vector $p(x, t)$ at time t, the probability that the system will be in the state x at time $t + dt$ is given by

$$p(x;t+dt) = p(x;t)(1 - \sum_{\mu=1}^{m} a_{\mu}(x)dt) + \sum_{\mu=1}^{m} p(x - v_{\mu};t)a_{\mu}(x - v_{\mu})dt \quad (1)$$

where $a_{\mu}(x)dt$ denotes the probability that the μ^{th} reaction will happen in a time step of length dt and v_{μ} is the stoichiometric transition vector. From Equation 1, the differential equation known as *Chemical Master Equation* (CME) (Gillespie D. T., 1977) can be derived:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\dot{p}(x;t) = -p(x;t)\sum_{\mu=1}^{m} a_{\mu}(x) + \sum_{\mu=1}^{m} p(x-\nu_{\mu};t)a_{\mu}(x-\nu_{\mu}) \tag{2}$$

By considering all the reactions beginning or ending at state x, the time derivative of the probability density of state x can be written in vector form as (Munsky & Khammash, 2006):

$$p(x;t) = [p(x;t).....p(x-\nu_m;t)]\begin{bmatrix} -\sum_{\mu=1}^{m} a_{\mu}(x) \\ ..... \\ a_m(x-\nu_m) \end{bmatrix} \tag{3}$$

As n is finite, the number of different possible vectors for x is countable. Thus, we can map all the possible x's $\in N^n$ to a sequence $x_1, x_2, \cdots$ of elements in N. Let us represent that enumeration as X:= [x_1, $x_2, \cdots$] and rewrite Equation 3 as

$$\dot{P}(X;t) = P(X;t)A \tag{4}$$

where $P(X; t)=[p(x_1, t),p(x_2, t)...]$ is the complete probability density state vector at time t and A is the state reaction matrix. The matrix A has the following properties: it is independent of t, all of its diagonal elements are non-positive, and all of its off-diagonal elements are non-negative and all the rows sum to exactly zero. For the case of finite number of reachable states, the solution can be computed as:

$$P(X;t) = P(X;0)e^{At} \tag{5}$$

An approach to reduce the complexity of simulation of the above solution using Kronecker product analysis and Zassenhaus formula have been considered in (Caglar & Pal, 2013).

CME Model Simulation

The complexity of CME models can be tremendous to allow for direct computation of the analytical solution of the probability density state vector. For instance, n = 5 molecular species with number of molecules ranging from 0 to 99 will have a total of $100^5=10^{10}$ states whose probabilities need to be computed at each time duration. Thus, Monte Carlos based techniques are often utilized to generate the time varying probability density state vectors. The most commonly used technique is known as Gillespie's Stochastic Simulation Algorithm (SSA) (Gillespie D. T., 1977).

The SSA approach is statistically exact in the sense that infinite number of simulations will produce the same distribution as the analytical solution of Equation 2 (Székely Jr. & Burragea, 2014).

The SSA approach simulates multiple trajectories based on the governing CME by selecting the time when the next reaction will happen and what the reaction will be.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

SSA for generating a trajectory can be described using the following algorithm (Higham 2008)

1. Evaluate $a_\mu(x)$ at time t for μ=1 to m.
2. Draw random numbers r_1 and r_2 based on uniform (0 1).
3. Reaction j is selected that satisfies

$$\frac{\sum_{\mu=1}^{j-1} a_\mu(x)}{\sum_{\mu=1}^{m} a_\mu(x)} \leq r_1 < \frac{\sum_{\mu=1}^{j} a_\mu(x)}{\sum_{\mu=1}^{m} a_\mu(x)}$$

4. Time step for next reaction

$$\tau = \frac{-\ln(r_2)}{\sum_{\mu=1}^{m} a_\mu(x)}$$

5. Update time $t=t+\tau$ and state $x=x+v_j$
6. Repeat steps 1-5 until stopping criteria is met.

The stopping criteria is often an upper limit on time or the number of iterations. Readers are referred to (Gillespie D. T., 1977; Higham, 2008) for further details on SSA algorithm. Software for simulation of SSA are also available such as COPASI (Hoops & et.al, 2006), STOCKS (Kierzek, 2002), BioNets (Adalsteinsson, McMillen, & Elston, 2004) and StochKit2 (Sanft, Wu, Roh, Fu, & Petzold, 2011).

The SSA algorithm is simple to understand and implement but has high computational complexity. Tau Leaping (Gillespie D. T., 2000) is an approximation approach to SSA that considers all reactions occurring in a time interval τ before updating the propensity functions $a_\mu(x)$.

The simulation of CME models can also be facilitated by avoiding inclusion of molecular states that will never be reached or will likely have infinitesimal probabilities. (Cao & Liang, 2008) presents an optimal algorithm for simulating the probability landscape for networks with small copy numbers that avoids attempting infeasible state transitions and provides a concise state transition matrix devoid of redundant information. (Munsky & Khammash, 2006) approximates the solution of CME model by truncating the state space to remove states with minimal probabilities and provides a bound on the difference between the truncated space solution and the true solution. (Cao, Lu, & Liang, 2010) considers a method termed discrete chemical master equation (dCME) that is able to directly compute the full steady state probability distribution of lysogeney maintenance network in phage lambda.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MODEL APPROXIMATIONS

CME to Stochastic Differential Equation

An approximation to the discrete stochastic model in the form of a stochastic DE known as *chemical Langevin Equation (CLE)* (Gillespie D. T., 2000) is often used to make the solution analytically and numerically tractable. The approximation can be represented by the following stochastic DE (Cai & Wang, 2007):

$$x(t+dt) = x(t) + \sum_{\mu=1}^{m} \nu_{\mu} a_{\mu}(x(t))dt + \sum_{\mu=1}^{m} \nu_{\mu} \sqrt{a_{\mu}(x(t))} N_{\mu}(t) \tag{5}$$

where $N_{\mu}(t)$ for $\mu=1,\ldots m$ are independent Gaussian random variables with zero mean and unit variance. The conditions for validity of the approximation require that an infinitesimal time interval *dt* is available under which 1) there is no significant change in the propensity functions and 2) the expected number of occurrences of each reaction in *dt* is significantly greater than one. As shown in (Gillespie D. T., 2000), the first two moments of CME and CLE can be related using the truncation of chemical Kramers-Moyal equations (CKME). CKME can be considered an approximation of CME based on Taylor Series expansion and truncation of CKME will result in chemical Fokker-Plank equation representing the CLE. For details, readers are referred to (Gillespie D. T., 2000).

CME to Differential Equation

The relationship between CME and deterministic Differential Equation (DE) models have been well studied analytically by multiple researchers primarily from the perspective of similarity of the average predictive behavior of the two models (Kurtz, 1972; Kampen, 1991). Since the topic of CME to DE approximations have been well studied, this section will illustrate the relationship from a very basic perspective with simulation examples. For detailed analysis, readers are referred to (Kurtz, 1972; Kampen, 1991).

Example

In this example, we will apply CME to DE approximation technique to a generic switch (similar to a genetic toggle switch model without the detailed gene regulation) where there are two competing proteins U_1 and U_2, each of which inhibits the transcription of the other. The decay reactions $U_1 \rightarrow \varnothing$ and $U_2 \rightarrow \varnothing$ have propensities $c\Psi_1$ and $c\Psi_2$ respectively. The transcription of new copies of the proteins are guided by the reactions $\varnothing \rightarrow U_1$ and $\varnothing \rightarrow U_2$ with propensities b/(b $+\Psi_2$) and b/(b $+\Psi_1$) respectively. Our starting Stochastic Master equation model has parameter values $b = 4$ and $c = 0.2$. The corresponding stochastic master equation is

$$p(x;t+dt) = p(x;t)(1 - \sum_{\mu=1}^{m} a_{\mu}(x)dt) + \sum_{\mu=1}^{m} p(x-\nu_{\mu};t)a_{\mu}(x-\nu_{\mu})dt$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

with $m = 4$ and $v_1 = [-1, 0]$, $v_2 = [1, 0]$, $v_3 = [0, -1]$, $v_4 = [0, 1]$, $a_1([x_1, x_2]) = cx_1$, $a_2([x_1, x_2]) = b/(b+x_2)$, $a_3([x_1, x_2]) = cx_2$, $a_4([x_1, x_2]) = b/(b+x_1)$.

Our corresponding differential equation approximation is:

$$\frac{dy_1(t)}{dt} = \frac{b}{b + y_2(t)} - cy_1(t)$$

$$\frac{dy_2(t)}{dt} = \frac{b}{b + y_1(t)} - cy_2(t)$$

where $y_1(t)$ and $y_2(t)$ approximates $E_{x1}(x_1(t))$ and $E_{x2}(x_2(t))$ respectively. Figure 2 shows

$E_{x1}(x_1(t))$ (red stars) and $y_1(t)$ (blue line) for a starting distribution of $p(X) = [1, 0, 0 \cdots, 0]$ i.e. the number of protein molecules for both protein is zero for the initial state. It can be seen from Figure 4 that the approximation is very close to the expectation of the actual solution. We can solve for the steady state value of y_1 and y_2 analytically based on the equations $dy_1/dt=0$ and $dy_2/dt=0$.

If the steady state values of y_1 and y_2 are represented by y_{1s} and y_{2s}, then

$$\frac{b}{b + y_{2s}} - cy_{1s} = \frac{b}{b + y_{1s}} - cy_{2s} = 0 \tag{10}$$

Solving the equation gives $y_{1s} = \dfrac{-b + \sqrt{b^2 + 4b / c}}{2}$ which produces $y_{1s} = 2.9$ for the specific parameters of $b = 4$ and $c = 0.2$. The steady state value of 2.9 is also noticeable in Figure 4 for the simulated differential equation. Figure 3 shows the steady state marginal probability distribution of protein 1 simulated using the stochastic master equation. The expectation of this distribution 3.05 is very close to the analytically calculated differential equation steady state 2.9.

To study the effect of the control policy designed on the simpler approximate differential equation model being applied to the original stochastic master equation, we consider a control problem of maintaining the average expression of protein 1 to be around the value *10*. The control input is assumed to decrease the decay of the proteins by decreasing the value of the parameter c to *0.01*. Thus we have two systems, the original with $b = 4$ and $c = 0.2$ and the other corresponding to when control is switched on with parameters $b = 4$ and $c = 0.01$. The response of protein 1 of the network corresponding to parameters $b = 4$ and $c = 0.01$ is shown in Figure 4. The red stars denote the expectation of the distribution of protein 1 at every 2 second interval corresponding to the simulation of the stochastic master equation while the blue dots represent the output of the approximate differential equation model. The similarity of both the curves shows the closeness of the approximation. The value of the steady state corresponding to the DE model can be calculated by solving Equation 10 with $b = 4$ and $c = 0.01$ producing a value of $y_{1s} = 18.1$. The expectation of the steady state distribution calculated from the simulated CME model produces a similar value of *19.1*. The steady state probability distribution of protein 1 corresponding to the CME model is plotted in Figure 5.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

*Figure 2. The average protein 1 expression simulated using stochastic master equation (red *) and Differential Equation approximation (blue dots) for the original genetic toggle switch with parameters b = 4 and c = 0.2*

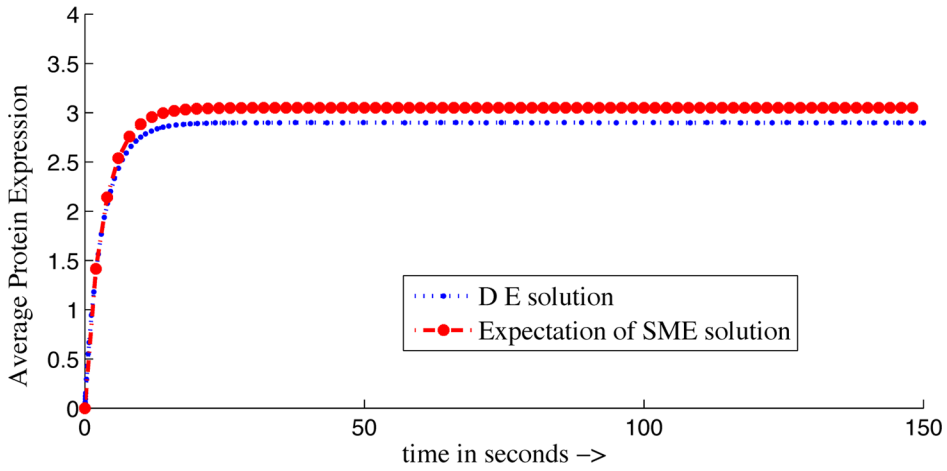


Figure 3. The steady state probability distribution of Protein 1 for the original genetic toggle switch

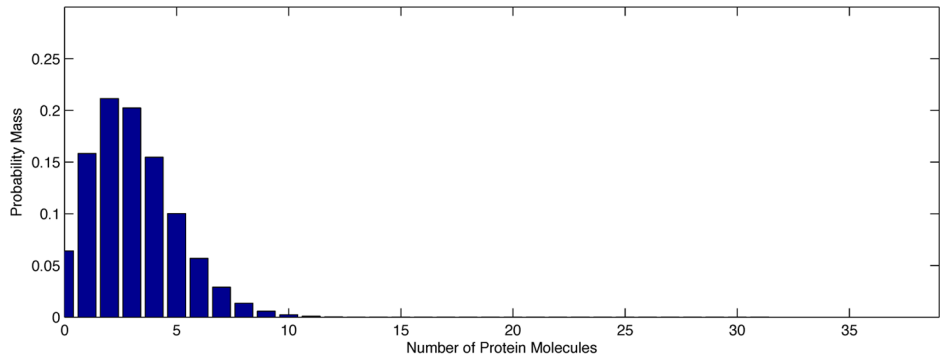
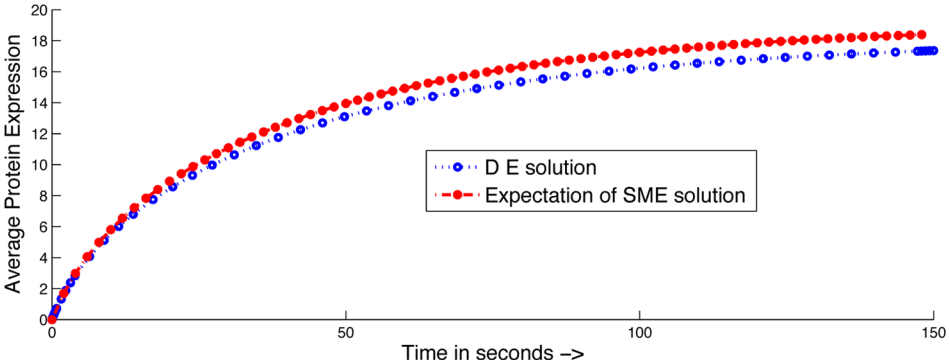
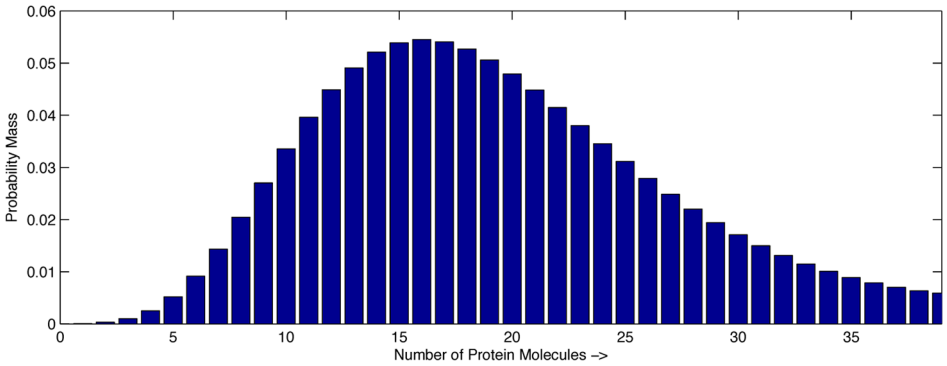


*Figure 4. The average protein expression simulated using stochastic master equation (red *) and Differential Equation approximation (blue dots) for the drug treated genetic toggle switch with parameters b = 4 and c = 0.01*



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

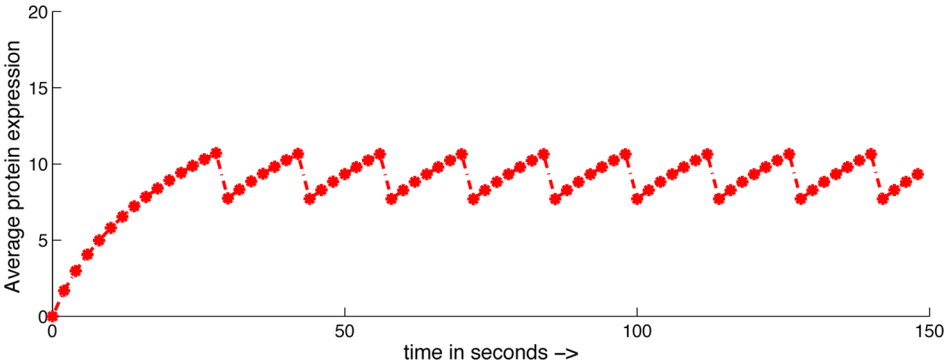
Figure 5. The steady state probability distribution of Protein 1 for the drug treated genetic toggle switch



As stated earlier, our control objective is to maintain the average expression of protein 1 to be around the value 10. We design a simple on off controller with deadzone based on the differential equation model which produces the following control policy: use system 1 (original system, b=4, c=0.2) if the average protein expression is above 10.5 and use system 2 (system when control is on, b=4, c=0.01) if the average protein expression is below 9.5. If the protein expression is between 9.5 and 10.5, maintain the earlier control policy. This is done to avoid indeterminate switching around the reference value of 10. The result of this control policy applied to the CME models is shown in Figure 6.

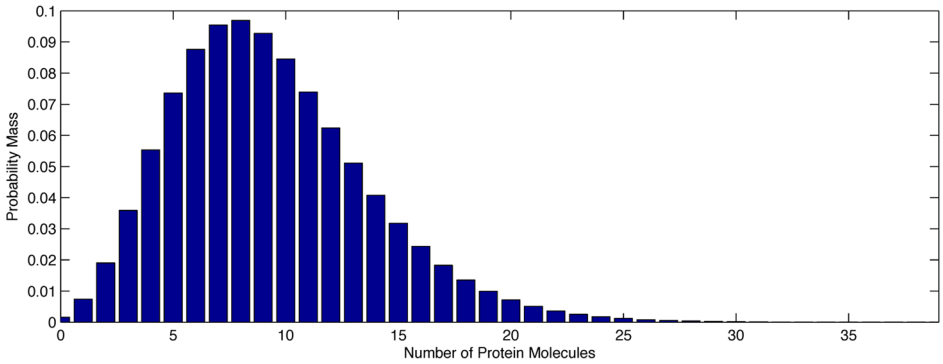
We notice that the control objective has been almost achieved with the average expression staying between 7.5 to 10.5. The variation could have been reduced if we apply control in much finer time interval than the current time interval of 2 seconds. In actual translational medicine, this time will be dependent on a number of factors like technical limitations on the sampling rate and application of control. Another way to approach the control can be by solving for the value of c producing the desired reference output and change the drug amount to achieve that specific value of c. In this case the drug is decreasing the degradation of the protein but not affecting the generation. We may be aware of the relationship between drug amount and parameter c but do not have the analytical solution relating the steady state output to the parameter c. We have to simulate the CME model for various values of c to arrive at the relationship between c and steady state expression. Our differential equation approximation

Figure 6. The average protein expression simulated using stochastic master equation after application of the on-off control policy designed using the differential equation approximation models



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. The steady state probability distribution of Protein 1 for the controlled genetic toggle switch

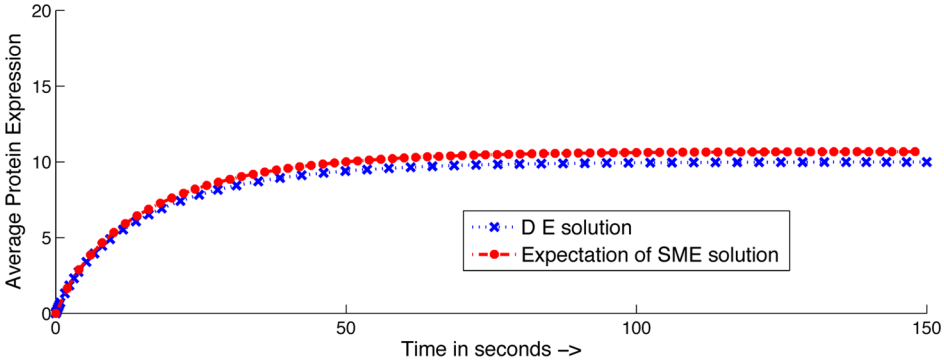


can analytically provide the value of *c* that produces the desired steady state output. Based on our desired response of 10, we get the required value of $c = \sqrt{\frac{b^2}{(20+b)^2 - b^2}}$ to be *0.0286*. Figure 8 shows the average protein expression simulated using stochastic master equation and differential equation model for protein 1 when the parameter *c* is maintained at *0.286*. The steady state output is very close to the desired output of *10*. In terms of numerical values, the expectation of the CME output for protein 1 at steady state is *10.69*.

CME to Probabilistic Boolean Networks

In an CME model, If we consider the number of possible x_i's to be finite, then the exact solution to the CME can be obtained through the exponential equation $P(X; t) = P(X, 0)e^{At}$. This kind of situation can arise in biological systems where the number of mRNA/Protein molecules that can be generated are bounded. If for systems, the dimension of X is infinite, we can apply the Finite State Projection approach (Munsky & Khammash, 2006) to arrive at a finite truncation of the state space. Henceforth,

*Figure 8. The average protein expression simulated using stochastic master equation (red *) and Differential Equation approximation (blue x) for the controlled genetic toggle switch with parameters b = 4 and c = 0.0286*



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

we will consider a CME with finite number of states M and our purpose is to transform that model to a smaller scale Probabilistic Boolean Network (PBN) model while maintaining some characteristics of the CME model.

The PBN model (Pal R., 2010) can be considered as a stochastic version of the Boolean Network (BN) (Kauffman S. A., 1993). A BN $B = (V, \boldsymbol{f})$ on n genes is defined by a set of nodes/genes $V = [x_1,..., x_n]$ and a vector of Boolean functions $\boldsymbol{f} = [f_1,..., f_n]$. The variable $x_i \in [0,1]$ represents the expression level of gene i, with 1 representing high and 0 representing low expression. The vector $\boldsymbol{f}$ represents the regulatory rules between genes. At every time step $t+1$, the value of x_i is predicted by the values of a set W_i of genes at the previous time step t, based on the regulatory function f_i, i.e. $x_i(t+1) = f_i(x_{i_1}(t),.....,x_{i_{k_i}}(t))$. The set of genes $W_i = \{x_{i_1},.....,x_{i_{k_i}}\}$ is called the predictor set of x_i, and the function f_i is called the predictor function of x_i. The pairs (x_i, W_i), $i = 1, ..., n$ induce a digraph G with edges $x_{i_j} \rightarrow x_i$ representing the structural dependencies among the genes. A state of B is a vector $\boldsymbol{s} = [x_1,..., x_n] \in [0, 1]^n$. All of the possible states of the Boolean network comprise its state space S which combined with the functions in $\boldsymbol{f}$ produces a digraph Γ called the state transition diagram of B. Γ represents the dynamics of the Boolean network and can be identified with a $2^n \times 2^n$ matrix P_n with rows and columns indexed by the states in B and entries $p_{ij} = 1$ or 0 if there is a transition from the state $s_i \rightarrow s_j$ in S.

Since constraints on the experimental data may limit the inference of a unique Boolean Network, the uncertainty in estimation along with the stochastic nature of genomic regulation is captured in the form of a PBN model where each gene can have multiple predictor functions. A *context-sensitive* PBN $A = A^{q,p}(V, \boldsymbol{F}, C)$ is defined by a set of nodes/genes $V = \{x_1,..., x_n\}$, a set of vector-valued Boolean functions $\boldsymbol{F} = \{\boldsymbol{f}_1,\ldots,\boldsymbol{f}_r\}$, $f_j: \{0,1\}^n \rightarrow \{0,1\}^n$, $j = 1,\ldots, r$ called *realizations* or *network functions*, a list of selection probabilities $C = \{c_1,\ldots, c_r\}$ for the corresponding realizations, the gene mutation/flipping probability p, and the realization switching probability q (Pal R., 2010). Updating the values of all genes in the network at time t is done synchronously according to the components of the currently used network function, and then the process is repeated. The choice of which network function $\boldsymbol{f}_j$ to apply is governed by a selection procedure. Specifically, at each time point t a random decision is made as to whether to switch the network function for the next transition, with a probability q of a switch being a system parameter. If a decision is made to switch the network function, then a new realization is chosen from among all of the possible realizations $\boldsymbol{f}_j \in \boldsymbol{F}$ of A, according to their individual selection probabilities $c_j \in C$. In other words, each network function $\boldsymbol{f}_j$ represents a deterministic BN B_j and the PBN behaves as a fixed BN until a random decision (with probability of q) is made to change the network function according to the probabilities $\{c_1,\ldots, c_r\}$ from among $\{\boldsymbol{f}_1,\ldots,\boldsymbol{f}_r\}$. In addition to the network switching and selection in the PBN model, there is mechanism which models random gene mutations, i.e. at each time point t there is a probability p of any gene changing its value uniformly randomly. Thus, the PBN model can account for the uncertainties in both data and model selection. The PBN A shares the same state space S with its realizations, and the state transition diagrams Γ_j of the individual B_j's combine naturally into a stochastic state transition diagram Γ representing the dynamics of A. As in the case of deterministic BNs Γ can be identified with a stochastic $2^n \times 2^n$ matrix P_n, also known as *transition* matrix, with non-negative entries p_{ij} and having the property $\sum_{j=1}^{n} p_{ij} = 1$, $i = 1,\ldots, n$. Using this matrix, the dynamics of A can be described using the well-developed theory of Markov chains. One should notice that

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

if the probability of gene flipping p is positive then the Markov chain representing the dynamics of the network is ergodic which implies that it possesses a *steady-state probability distribution* π.

Both the CME and PBN models belong to the class of discrete state stochastic dynamic models but differ in the level of discretization (coarse scale for PBN vs. fine-scale for CME) and the discrete (PBNs) or continuous time (CMEs) updating of the states. In this section, our overall goal is to study mappings from CMEs to PBNs that can maintain some desirable properties. As the steady state distribution of the genetic regulatory network is considered representative of the phenotype, our objective will be to maintain the steady state distribution. The finite state continuous time CME model can be approximated by a discrete-time sampled Markov chain by using a suitable time period Δt. The M × M state transition probabilities of the discrete-time Markov Chain denoted by P_Δ is given by the equation $P_\Delta = e^{A\Delta t}$. We will map these transition probabilities to an N × N transition probability matrix of a PBN denoted by P_r where N will be 2^n for binary levels and 3^n for the case of ternary levels. We will consider binary PBNs from now onwards but the derivation for more than 2 levels will be similar. Let T_i for $i \in [1, 2, \cdots n]$ denote the thresholds for binarization. For instance, if $x(t)=[\Psi_1,\Psi_2,\ldots,\Psi_n]$, then the corresponding state for the PBN is $y=[\psi_1,\psi_2,\ldots,\psi_n]$ where $\psi_1=1$ if $\Psi_i \geq T_i$ or $\psi_i=0$ if $\Psi_i < T_i$ for all $i \in [1, 2, \cdots n]$. As the states of the PBN will be n-dimensional vectors ranging from 00..0 to 11...1, we can represent them by the decimal equivalent state z_t where z_t ranges from 1 to 2^n. Let us consider a sequence $0=a_0,a_1,\ldots,a_N=M$ such that $a_i - a_{i-1}$ for $i \in \{1, 2, ..., 2^n\}$ denote the number of states in the CME model that map to state i in the PBN model. The states of the CME model will be denoted by w_t ranging from 1 to M. The states will be re-ordered such that the states $a_{i-1} + 1$ to state a_i of the CME will be mapped to state i of the PBN for $i = 1, ..., N$. Note that the subscript Δ refers to the higher dimensional discrete time CME model whereas the subscript r refers to the lower dimensional PBN model.

Let the state of the PBN at time be $z_t=i$ and the state at time $t+1$ be $z_{t+1}=j$. The transition probability of going from state i to state j is given by

$$
\begin{aligned}
P_r(i,j) &= P(z_{t+1}=j \mid z_t=i) = \sum_{j_1=a_{j-1}+1}^{a_j} P(\omega_{t+1}=j_1, z_t=i) \\
&= \frac{\sum_{j_1=a_{j-1}+1}^{a_j} P(\omega_{t+1}=j_1, z_t=i)}{P(z_t=i)} = \frac{\sum_{j_1=a_{j-1}+1}^{a_j}\sum_{i_1=a_{i-1}+1}^{a_i} P(\omega_{t+1}=j_1, \omega_t=i_1)}{\sum_{i_2=a_{i-1}+1}^{a_i} P(\omega_t=i_2)} \\
&= \frac{\sum_{j_1=a_{j-1}+1}^{a_j}\sum_{i_1=a_{i-1}+1}^{a_i} P(\omega_{t+1}=j_1 \mid \omega_t=i_1)P(\omega_t=i_1)}{\sum_{i_2=a_{i-1}+1}^{a_i} P(\omega_t=i_2)} = \frac{\sum_{j_1=a_{j-1}+1}^{a_j}\sum_{i_1=a_{i-1}+1}^{a_i} P_\Delta(i_1,j_1)P(\omega_t=i_1)}{\sum_{i_2=a_{i-1}+1}^{a_i} P(\omega_t=i_2)}
\end{aligned}
\tag{11}
$$

We next show that if we use the steady state probabilities of the CME model for $P(\omega_t=i_1)$ in Equation 11, then the steady state probabilities of the resultant PBN, P_r will be the same as the collapsed steady state probabilities of the CME model. Here collapsing refers to the summation of the state probabilities

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of the detailed model that map to a certain state of the PBN model. For instance, if η represents the M dimensional steady state probability vector for the CME model, then the collapsed steady state vector will be an N dimensional vector represented by

$$\zeta(i) = \sum_{c=a_{i-1}+1}^{a_i} \eta(c) \text{ for } i=1,\ldots N.$$

We next prove the equivalence of ζ and the steady state probabilities of the PBN represented by the following state transition probabilities:

$$P_r(i,j) = \frac{\sum_{j_1=a_{j-1}+1}^{a_j} \sum_{i_1=a_{i-1}+1}^{a_i} P_\Delta(i_1,j_1)\eta(i_1)}{\sum_{i_2=a_{i-1}+1}^{a_i} \eta(i_2)} \tag{12}$$

Theorem I: (Pal & Bhattacharya, 2010) *With Equation 12 denoting the transition probabilities of the PBN, the steady state probability distribution vector, π,* o*f the PBN is given by*

$$\pi(i) = \sum_{i_2=a_{i-1}+1}^{a_i} \eta(i_2) = \zeta(i)$$

Proof: Assuming an ergodic Markov Chain, ζ is the steady state distribution of the PBN if it satisfies the equation $\sum_{i=1}^{N} \zeta(i) P_r(i,j) = \zeta(j)$ for each $j \in 1,2,\ldots,N$

Since η is the steady state distribution for the $M \times M$ CME model P_Δ, it satisfies $\eta(j_1) = \sum_{i_0=1}^{M} \eta(i_0) P_\Delta(i_0,j_1)$.

Thus

$$\zeta(j) = \sum_{j_1=a_{i-1}+1}^{a_j} \eta(j_1) = \sum_{j_1=a_{i-1}+1}^{a_j} \sum_{i_0=1}^{M} \eta(i_0) P_\Delta(i_0,j_1) =$$

$$\sum_{j_1=a_{i-1}+1}^{a_j} \sum_{i=1}^{N} \sum_{i_1=a_{i-1}+1}^{a_i} \eta(i_1) P_\Delta(i_1,j_1) = \sum_{i=1}^{N} \sum_{i_2=a_{i-1}+1}^{a_i} \eta(i_2) P_r(i,j) = \sum_{i=1}^{N} P_r(i,j)\zeta(i)$$

This shows that ζ satisfies the steady state probability equations of the PBN and is equal to π assuming an ergodic Markov Chain. We have established that if the transition probabilities of the PBN are decided based on Equation 12, then the steady state probabilities of the generated PBN is same as the collapsed steady state probabilities of the original detailed network. Thus, we are able to maintain the collapsed

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

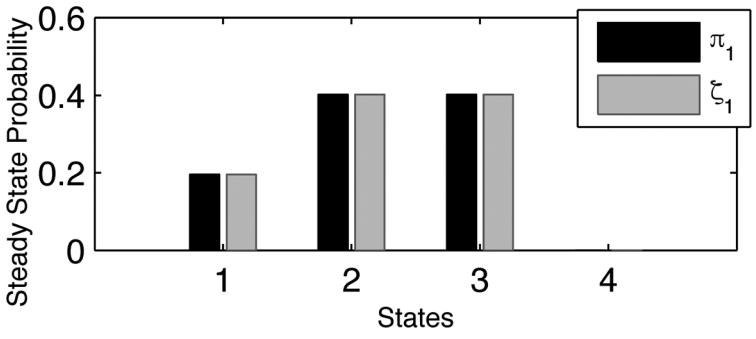
steady-state probabilities, which might be indicative of a phenotype. If we revisit Equation 12, we notice that $P_r(i,j)$ can be calculated from the transitions of the network once it has reached the steady state. If we observe the steady-state transitions of the genetic regulatory network modeled by the CME for a long time and generate the state-transition probabilities of the PBN $P_r(i,j)$ based on the number of transitions occurring from state i to state j divided by the total number of transitions from state i, we will arrive at Equation 12. Thus, the mapping described by Equation 12 exists between a network represented by a CME model and a coarse-scale PBN model when the transition probabilities of the PBN are inferred based on state transition data at steady-state.

The following two examples from (Pal & Bhattacharya, 2010) consider well-studied stochastic master equation models of biological networks for illustrating the maintenance of collapsed steady state behavior when the mapping is based on Equation 12.

Example 1

We apply the CME to PBN approximation to the same genetic toggle switch model presented in the earlier case of CME to DE approximation. We have two competing proteins U_1 and U_2, each of which inhibits the transcription of the other. The decay reactions $U_1 \rightarrow \varnothing$ and $U_2 \rightarrow \varnothing$ have propensities $c\psi_1$ and $c\psi_2$ respectively. The transcription of new copies of the proteins are guided by the reactions $\varnothing \rightarrow U_1$ and $\varnothing \rightarrow U_2$ with propensities b/(b $+\psi_2$) and b/(b $+\psi_1$) respectively. Our starting Stochastic Master equation model has parameter values b = 4 and c = 0.05. We truncate the maximum number of protein molecules for U_1 and U_2 to 39 resulting in a total number of 40x40 = 1600 possible states for the CME. Based on the propensity functions and parameter values b=0.4 and c=0.05, the state reaction matrix A_1, is generated and the continuous Markov chain is approximated by a discrete chain using the following equation $P_{\Delta_1} = e^{A_1 \Delta t}$ with a time step of Δt=0.1 seconds. The thresholds for binarization of the proteins U_1 and U_2 are selected to be T_1=T_2=6 molecules. After binarization, the two protein reduced network has 2x2=4 states. We calculated the transition probabilities of the reduced model (PBN) using Equation 12. In our case, the Markov chain is ergodic as all the states are communicating, recurrent and aperiodic. Due to the ergodicity, P_{Δ_1} has a unique steady state probability distribution. The equivalence of the steady state probability distribution π_1 of P_{r_1} and the collapsed steady state probability distribution ζ_1 of P_{Δ_1} is shown in Figure 9 (Pal & Bhattacharya, 2010).

Figure 9. Steady State Probability Distributions π_1 and ζ_1
Pal and Bhattacharya (2010) © 2010 IEEE



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Example 2

In this example, we apply the CME to PBN mapping procedure to a model representing the life cycle of bacteriophage-λ (Hegland, Burden, Santoso, MacNamara, & Booth, 2007) which involves five proteins. The decay reactions $U_i \rightarrow \varnothing$ for i=1,2,…,5 are with propensities $c_i \Psi_i$ for i=1,2,…,5. The generation reactions with their respective propensities are as follows:

1. $\varnothing \rightarrow U_1$ with propensity $\dfrac{a_1 b_1}{b_1 + \Psi_2}$
2. $\varnothing \rightarrow U_2$ with propensity $\dfrac{(a_2 + \Psi_5) b_1}{b_1 + \Psi_1}$
3. $\varnothing \rightarrow U_3$ with propensity $\dfrac{a_3 b_3 \Psi_2}{b_3 \Psi_2 + 1}$
4. $\varnothing \rightarrow U_4$ with propensity $\dfrac{a_4 b_4 \Psi_3}{b_4 \Psi_3 + 1}$
5. $\varnothing \rightarrow U_5$ with propensity $\dfrac{a_5 b_5 \Psi_3}{b_5 \Psi_3 + 1}$

The generation reactions for the proteins originate from the mechanism of transitions between two phases, lysogenic and lytic, present typically in the life of a bacteriophage-λ. Alterations from one phase to another are regulated by a toggle switch of two competing proteins U_1 and U_2, which inhibit each other's transcription. In the lysogenic phase U_2 is abundant while suppressing the production of U_1 via the reaction 1. The generation of U_2 via reaction 2, on the other hand, depends on the amount of U_5. Under a stressed situation U_5 decreases as the phage transitions into the lytic stage. The parameters a_i, b_i, c_i are taken from (Hegland, Burden, Santoso, MacNamara, & Booth, 2007) and shown in Table 1.

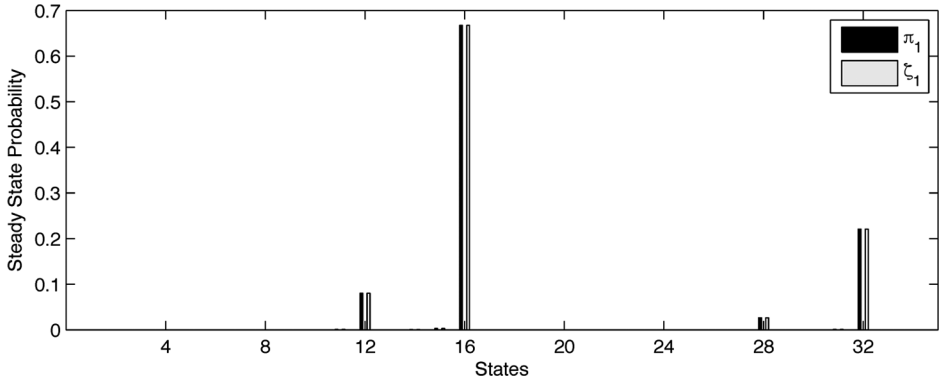
The maximum number of protein molecules for U_1, U_2, U_3, U_4, U_5 are truncated to be at 3, 24, 3, 3, and 2, respectively. Thus, the total number of possible states for the CME is 4x25x4x4x3=4800. Using the above propensity functions and values for parameters shown in Table I, the state reaction matrix A_1 is computed and the continuous time Markov chain is approximated by a discrete time chain using equation $P_{\Delta_1} = e^{A_1 \Delta t}$ with a time step of Δt=0.1 seconds. The binarization of the proteins is achieved using thresholds 2, 8, 2, 1, and 1, respectively. After binarization, the five protein reduced network has 2^5=32 states. The transition probabilities of the reduced model (PBN) are calculated using Equation 12. Since all the states are communicating, recurrent and aperiodic, the Markov chain P_{Δ_1} is ergodic and hence

Table 1. Parameters for the example II CME model

	i=1	i=2	i=3	i=4	i=5
a_i	0.5	1	0.15	0.3	0.3
b_i	0.12	0.6	1	1	1
c_i	0.0025	0.0007	0.0231	0.01	0.01

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Steady State Probability Distributions π_1 and ζ_1 for example II
Pal and Bhattacharya (2010) © 2010 IEEE



P_{Δ_1} has a unique steady state probability distribution. The equivalence of the steady state probability distribution π_1 of P_{r1} and the collapsed steady state probability distribution ζ_1 of P_{Δ_1} is shown in Figure 10.

Transient Behavior of Coarse-Scale CME Approximation

The mapping presented in Equation 12 maintains the collapsed steady state behavior without enforcing any constraints on the collapsed transient dynamics. Thus, in this section we will evaluate the deviation of the transient dynamics of the PBN as compared to the collapsed CME state probability distribution. Let us consider the difference equation $d_j(t+\Delta t) = |\gamma j_(t+\Delta t) - \tau j(_t+\Delta t)|$ where $\gamma j(_t)$ denote the collapsed probability of states aj–$_{l+1}$ to aj $_o$f the CME model at any time t *a*nd $\tau j(_t)$ denote the probability of state j *a*t any time t *f*or the PBN model. We can bound the difference equation dj$(_t+\Delta t)$ for $\gamma j(_t)=\tau j(_t)$ in the following way (Pal & Bhattacharya, 2012; Pal, Bhattacharya, & Caglar, 2012):

$$d_j(t+\Delta t) \leq \max_{i_1 \in [1,\ldots,N]} \left(\max_{i_2 \in S_{i_1}} q(i_2, j) - \min_{i_2 \in S_{i_1}} q(i_2, j) \right)$$

where $q(i_2, j) = \sum_{j_1 = a_{j-1}+1}^{a_j} P_{\Delta}(i_2, j_1)$.

Similarly, we can extend the difference equation for multiple time steps as follows (Pal & Bhattacharya, 2012):

$$d_j(T\Delta t) \leq \beta \left(\frac{1-\mu_m}{1-\mu_m} \right)^T$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

where $\beta = \max_{j\in[1,..N]} \beta_j$ and $\beta_j = \max_{i_1\in[1,...,N]} (\max_{i_2\in S_{i_1}} q(i_2,j) - \min_{i_2\in S_{i_1}} q(i_2,j))$ and $\mu_m = \max_{i_1\in[1,...,N]} (\sum_{i_1=1}^{N} | R(i_1,j) - \rho |)$ where ρ is any constant and $R(i_1,j) = \min_{i_2\in[a_{i_1-1}+1,.....,a_{i_1}]} q(i_2,j)$.

The above-mentioned bounds for single and multiple time steps provide the worst-case bounds that can be difficult to compute as they require the detailed knowledge of the fine-scale transition probability matrix. Thus, a possible approach is to construct stochastic bounds or bounds on the expected deviation based on possible distributions of the entries of the transition probability matrix. Multiple such bounds based on the uniform distribution or distribution centered around different means for the entries of the state transition probability matrix have been considered in (Pal & Bhattacharya, 2012). The next two theorems from (Pal & Bhattacharya, 2012) provide the stochastic bounds for single and multiple steps. The proofs are available in (Pal & Bhattacharya, 2012).

Theorem II (Pal & Bhattacharya, 2012) The expected bound on the deviation β_j for single time step is given by the following equation:

$$E(\beta_j) < 2\sigma(2\log L - \log\log L)^{\frac{1}{2}}$$

where σ denotes the standard deviation for the distribution of $q(i,j)$ and $L = \max_{i=1,...,N} | a_i - a_{i-1} |$.

Theorem III (Pal & Bhattacharya, 2012)

$$E(\mu_m) < (N-1)\sqrt{\frac{2}{\pi}}\sigma_1 + (\xi_2 - \xi_1) erf(\frac{\xi_2 - \xi_1}{\sqrt{2}\sigma_2}) + \sqrt{\frac{2}{\pi}}\sigma_2 e^{-(\xi_2-\xi_1)}$$

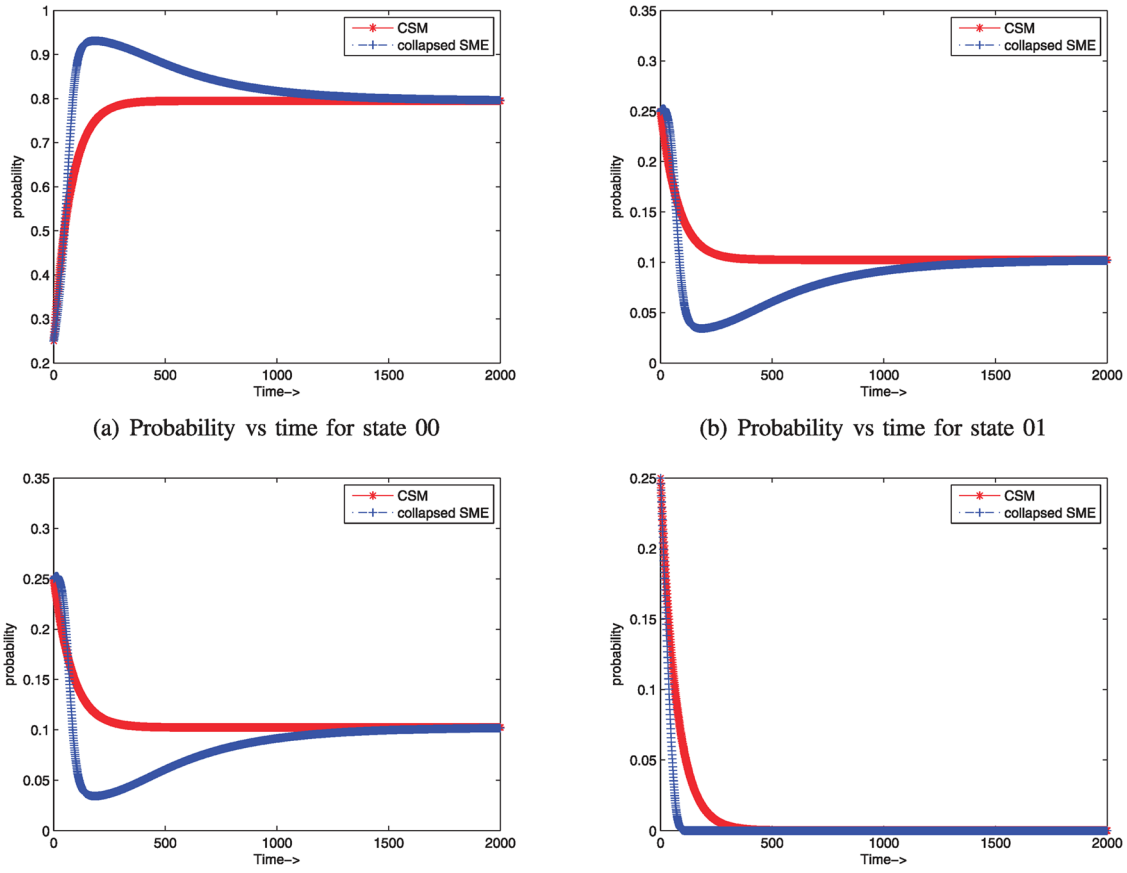
where $erf(x) = \frac{2}{\sqrt{\pi}} \int_{t=0}^{x} e^{-t^2} dt$ denotes the error function and the distribution of $R(i,j) = \min_{i_2\in[a_{i-1}+1,...,a_i]} q(i_2,j)$ is approximated by $\aleph(\xi_1,{\sigma_1}^2)$ and $\aleph(\xi_2,{\sigma_2}^2)$ for $i \neq j$ and $i = j$ respectively.

Example 3

In this example, we consider the earlier considered Genetic Toggle Switch in example I with parameters b = 0.8 and c = 0.035 as the initial system. We truncate the maximum number of protein molecules for U_1 and U_2 to 39 resulting in a total number of 40x40 = 1600 possible states for the CME. Based on the propensity functions and parameter values b=0.8 and c=0.035, the state reaction matrix A_1, is generated and the continuous Markov chain is approximated by a discrete chain using the following equation $P_{\Delta_1} = e^{A_1 \Delta t}$ with a time step of Δt=0.5 seconds. The thresholds for binarization of the proteins U_1 and U_2 are selected to be T_1=T_2=20 molecules. After binarization, the two protein reduced network has 2x2=4 states. We calculated the transition probabilities of the reduced model (PBN) using Equation 12. The time evolution of the probability of the four states starting from a uniform distribution is shown in Figure 11.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Time evolution of probability starting from a uniform distribution for the genetic switch with parameters b = 0.8 and c = 0.035. Red denotes the coarse-scale model evolution and blue + denotes the evolution of the collapsed CME model.
Pal and Bhattacharya (2012) © 2012 IEEE



(a) Probability vs time for state 00

(b) Probability vs time for state 01

(c) Probability vs time for state 10

(d) Probability vs time for state 11

The actual deviation characteristics of the PBN probability and the collapsed CME model are shown in Table II. State denotes the four states (00; 01; 10 and 11) of the reduced model. D_s and D_m denotes the actual maximum deviation in a single time step and multiple time steps, respectively. UB_s and UB_m denotes the stochastic bounds for single time step and multiple time steps, respectively. The ε used to calculate the stochastic bound is $2\frac{(1-d)}{(M-1)}$ with d=0.4 where d is an approximation of a diagonal entry. We are assuming that each diagonal entry is greater than 0.4 which is reasonable as we have a small δ$t$of 0.25 seconds. The μ_m used to calculate the multiple step bound is estimated as follows: the transition probability matrix is given by $e^{A\delta t}$ where A is the reaction matrix. We approximate the transition probability matrix by I+$A\delta t$ $u$sing the first two terms of the series expansion. Based on this approximation, we generated the approximate transition probability matrix and calculated μ s from the approximate matrix. For our case, δt=0.25, c=0.0035, b=0.8, and t1=$_t$2=$_2$0. To calculate the multistep bound, we

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Transient Response Bounds for the Genetic Switch Model with Parameters b= 0.8 and c = 0.035 (Pal & Bhattacharya, 2012) © 2012 IEEE

State	Ds	Dm	UBs	UBm
00	0.0054	0.2257	0.0479	0.2596
01	0.0027	0.0895	0.0479	0.2596
10	0.0027	0.0917	0.0479	0.2596
11	0.0016	0.0784	0.0479	0.2596

used the maximum of μ s 0.8154 as the μ=m. From Table 2, we note that the stochastic bound for multi-steps (0.2596) is very close to the maximum actual deviations for each state (0.2257, 0.0895, 0.0917, 0.0784).

DISCUSSIONS

In this chapter, we discussed the broad classes of genetic regulatory network models and presented equivalence relationships concerning some of them. We primarily discussed the relationships between CME and PBN type of models that are both discrete state stochastic models with Markovian property but with a hugely different size of the state space. Equivalence relationships among other model types have also been explored in the literature, for instance (Davidich & Bornholdt, 2008) shows that the Boolean Network model is a coarse-grain limit of the differential equation model and demonstrates it on the budding yeast cell cycle network. (Lähdesmäki, Hautaniemi, Shmulevich, & Yli-Harjaa, 2006) illustrates that probabilistic Boolean networks and dynamic Bayesian network models can represent the same multivariate state probability distribution over their common variables. Based on our modeling objectives and available data, different model types can be suitable for different scenarios. For instance, modeling of the basic average inhibition and up regulation of cancer pathways using targeted drugs can often be modeled using Boolean Network based models (Grasso, et al., 2015; Layek, Datta, & Dougherty, 2011; Pal & Berlow, 2012) . However, modeling of detailed stochasticity of gene expression (Kærn, et al. 2005) will require fine scale CME or stochastic differential equation based models. Furthermore, modeling of tumor heterogeneity will either require separate deterministic models for individual sub types or a stochastic model that incorporates the behavior of the diverse subtypes. The most commonly used model of ordinary differential equations is valid when the number of molecules of the modelled entities is sufficiently large. In terms of CME propensity functions, the CME to DE approximation is valid when the propensity functions are linear relationships of state variables (Szallasi, Stelling, & Periwal, 2006).

Furthermore, the mathematical model selected for representing a biological phenomenon will also be dependent on the data type used for inference of the model parameters. For instance, microarray data will provide the average expression of a collection of cells and cannot be used to infer fine-scale stochastic model parameters but are rather suited for DE or coarse scale Boolean Network kind of models. For estimating the parameters of a fine scale stochastic model, we would require individual cell measurements using techniques such as green fluorescent proteins or single cell sequencing. The sampling time period used in data measurements will also constraint the inference of model parameters. For instance, modeling a cyclic biological activity such as circadian rhythm will require measurement time step to be

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

smaller than the period of the biological activity. Some studies measure the gene or protein expression of a cell line at a fixed time and aggregate the responses across multiple samples to reflect the system behavior. This approach can work when the sample behaviors are similar and we have limited variations among the cell lines. However, for inference of dynamic model parameters, time series data is highly desirable as compared to steady state data. The model inference process is often independent of the experimental data generation process and majority of inference analysis is conducted without access to experimental data that is most suitable for high accuracy estimation of model parameters. Model based experiment design can improve the accuracy and decrease uncertainty in inference of model parameters.

One of the benefits of exploring the relationships between various models of GRNs lies in its potential to be used in building whole cell mathematical models where different model types represent the sub-networks. There have been recent attempts at creating whole cell models of smaller organisms such as (Karr, et al., 2012) that creates a whole cell computational model of the life cycle of the human pathogen *Mycoplasma genitalium* by combining 28 dissimilar submodels representing various levels and processes in the cells. (Covert, Xiao, Chen, & Karr, 2008) considers the specific case of integrating flux-balance based central carbon metabolic and transcriptional regulatory model with an ordinary differential equation based detailed model of carbohydrate uptake control for Escherichia coli. (Chandrasekaran & Price, 2010) considers the integration of transcriptional regulatory network with corresponding metabolic network for Escherichia coli and Mycobacterium tuberculosis. Thus, the detailed understanding of relationships between various models of GRNs will enable us to effectively integrate existing sub-models for an organism that have been designed using different model types.

REFERENCES

Adalsteinsson, D., McMillen, D., & Elston, T. C. (2004). Biochemical Network Stochastic Simulator (BioNetS): Software for stochastic modeling of biochemical networks. *BMC Bioinformatics*, *5*(1), 24. doi:10.1186/1471-2105-5-24 PMID:15113411

Arkin, A., Ross, J., & McAdams, H. H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics*, *149*(4), 1633–1648. PMID:9691025

Berlow, N., Davis, L., Cantor, E., Seguin, B., Keller, C., & Pal, R. (2013). A new approach for prediction of tumor sensitivity to targeted drugs based on functional data. *BMC Bioinformatics*, *14*(1), 239. doi:10.1186/1471-2105-14-239 PMID:23890326

Bishop, L. M., & Qian, H. (2010). Stochastic Bistability and Bifurcation in a Mesoscopic Signaling System with Autocatalytic Kinase. *Biophysical Journal*, *98*(1), 1–11. doi:10.1016/j.bpj.2009.09.055 PMID:20074511

Caglar, M. U., & Pal, R. (2013). Stochastic Model Simulation Using Kronecker Product Analysis and Zassenhaus Formula Approximation. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *10*(5), 1125–1136. doi:10.1109/TCBB.2013.34 PMID:24384703

Cai, X., & Wang, X. (2007). Stochastic Modeling and Simulation of Gene Networks. *IEEE Signal Processing Magazine*, *24*(1), 27–36. doi:10.1109/MSP.2007.273051

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cao, Y., & Liang, J. (2008). Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. *BMC Systems Biology*, *2*(30). doi:10.1186/1752-0509-2-30 PMID:18373871

Cao, Y., Lu, H. M., & Liang, J. (2010). Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18445–18450. doi:10.1073/pnas.1001455107 PMID:20937911

Chandrasekaran, S., & Price, N. D. (2010). Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in Escherichia coli and Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(41), 17845–17850. doi:10.1073/pnas.1005139107 PMID:20876091

Covert, M. W., Xiao, N., Chen, T. J., & Karr, J. R. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli. *Bioinformatics (Oxford, England)*, *24*(18), 2044–2050. doi:10.1093/bioinformatics/btn352 PMID:18621757

Davidich, M. I., & Bornholdt, S. (2008). Boolean Network Model Predicts Cell Cycle Sequence of Fission Yeast. *PLoS ONE, 3*(2), e1672.

De Jong, H. (2001). Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *Journal of Computational Biology*, *9*(1), 67–103. doi:10.1089/10665270252833208 PMID:11911796

Gillespie, D. T. (1977). Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Physical Chemistry*, *81*(25), 2340–2361. doi:10.1021/j100540a008

Gillespie, D. T. (2000). The chemical Langevin equation. *The Journal of Chemical Physics*, *113*(1), 297–306. doi:10.1063/1.481811

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, *115*(4), 1716–1733. doi:10.1063/1.1378322

Grasso, C., Tang, Y., Truffaux, N., Berlow, N. E., Liu, L., Debily, M. A., ... Monje, M. (2015). Functionally defined therapeutic targets in diffuse intrinsic pontine glioma. *Nature Medicine*, *21*(6), 555–559. doi:10.1038/nm.3855 PMID:25939062

Hegland, M., Burden, C., Santoso, L., MacNamara, S., & Booth, H. (2007). A solver for the stochastic master equation applied to gene regulatory networks. *Journal of Computational and Applied Mathematics*, *205*(2), 708–724. doi:10.1016/j.cam.2006.02.053

Higham, D. J. (2008). Modeling and Simulating Chemical Reactions. *SIAM Review*, *50*(2), 347–368. doi:10.1137/060666457

Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., ... Kummer, U. (2006). COPASI—a COmplex PAthway SImulator. *Bioinformatics (Oxford, England)*, *22*(24), 3067–3074. doi:10.1093/bioinformatics/btl485 PMID:17032683

Kærn, M., Elston, T. C., Blake, W. J., & Collins, J. J. (2005). Stochasticity in gene expression: From theories to phenotypes. *Nature Reviews. Genetics*, *6*(6), 451–464. doi:10.1038/nrg1615 PMID:15883588

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kampen, N. G. (1991). *Stochastic processes in physics and chemistry*. North Holland.

Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M. Jr, Bolival, B. Jr, ... Covert, M. W. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, *150*(2), 389–401. doi:10.1016/j.cell.2012.05.044 PMID:22817898

Kauffman, S. (1969). Homeostasis and Differentiation in Random Genetic Control Networks. *Nature*, *224*(5215), 177–178. doi:10.1038/224177a0 PMID:5343519

Kauffman, S. A. (1993). *The Origins of Order*. Oxford, UK: Oxford University Press.

Kierzek, A. M. (2002). STOCKS: STOChastic Kinetic Simulations of biochemical systems with Gillespie algorithm. *Bioinformatics (Oxford, England)*, *18*(3), 470–481. doi:10.1093/bioinformatics/18.3.470 PMID:11934747

Kurtz, T. G. (1972). The Relationship between Stochastic and Deterministic Models for Chemical Reactions. *The Journal of Chemical Physics*, *57*(7), 2976–2978. doi:10.1063/1.1678692

Lähdesmäki, H., Hautaniemi, S., Shmulevich, I., & Yli-Harjaa, O. (2006). Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks. *Signal Processing*, *86*(4), 814–834. doi:10.1016/j.sigpro.2005.06.008 PMID:17415411

Layek, R., Datta, A., & Dougherty, E. (2011). From biological pathways to regulatory networks. *Molecular BioSystems*, *7*(3), 843–851. doi:10.1039/C0MB00263A PMID:21161088

Munsky, B., & Khammash, M. (2006). The finite state projection algorithm for the solution of the chemical master equation. *The Journal of Chemical Physics*, *124*(4), 044104. doi:10.1063/1.2145882 PMID:16460146

Murphy, K. (2002). *Dynamic Bayesian Networks: Representation, Inference and Learning*. Berkeley, CA: University of California.

Pal, R. (2010). Context-sensitive probabilistic Boolean networks: Steady-state properties, reduction, and steady-state approximation. *IEEE Transactions on Signal Processing*, *58*(2), 879–890. doi:10.1109/TSP.2009.2030832

Pal, R., & Berlow, N. (2012). *A kinase inhibition map approach for tumor sensitivity prediction and combination therapy design for targeted drugs. In Pac Symp Biocomput* (pp. 351–362). Kohala Coast: World Scientific.

Pal, R., & Bhattacharya, S. (2010). Characterizing the Effect of Coarse-Scale PBN Modeling on Dynamics and Intervention Performance of Genetic Regulatory Networks Represented by Stochastic Master Equation Models. *IEEE Transactions on Signal Processing*, *58*(6), 3341–3351. doi:10.1109/TSP.2010.2043970

Pal, R., & Bhattacharya, S. (2012). Transient dynamics of reduced-order models of genetic regulatory networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *9*(4), 1230–1244. doi:10.1109/TCBB.2012.37 PMID:22411891

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Pal, R., Bhattacharya, S., & Caglar, M. (2012). Robust approaches for genetic regulatory network modeling and intervention: A review of recent advances. *IEEE Signal Processing Magazine*, *1*(29), 66–76. doi:10.1109/MSP.2011.942818

Sanft, K. R., Wu, S., Roh, M., Fu, J., & Petzold, L. R. (2011). StochKit2: Software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics (Oxford, England)*, *27*(17), 2457–2458. doi:10.1093/bioinformatics/btr401 PMID:21727139

Shea, M., & Ackers, G. (1985). The OR control system of bacteriophage lambda: A physical-chemical model for gene regulation. *Journal of Molecular Biology*, *181*(2), 211–230. doi:10.1016/0022-2836(85)90086-5 PMID:3157005

Szallasi, Z., Stelling, J., & Periwal, V. (2006). *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*. Cambridge, MA: The MIT Press. doi:10.7551/mitpress/9780262195485.001.0001

Székely, T. Jr, & Burragea, K. (2014). Stochastic simulation in systems biology. *Computational and Structural Biotechnology Journal*, *12*(20-21), 14–25. doi:10.1016/j.csbj.2014.10.003 PMID:25505503

This research was previously published in Emerging Research in the Analysis and Modeling of Gene Regulatory Networks edited by Ivan V. Ivanov, Xiaoning Qian, and Ranadip Pal; pages 52-79; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 13
Integral Equation Formulations and Related Numerical Solution Methods in Some Biomedical Applications of Electromagnetic Fields:
Transcranial Magnetic Stimulation (TMS), Nerve Fiber Stimulation

Dragan Poljak
University of Split, Croatia

Zoran Đogaš
University of Split, Croatia

Mario Cvetković
University of Split, Croatia

Maja Rogić Vidaković
University of Split, Croatia

Vicko Dorić
University of Split, Croatia

Jens Haueisen
Technical University Ilmenau, Germany

Ivana Zulim
University of Split, Croatia

Khalil El Khamlichi Drissi
Blaise Pascal University, France

ABSTRACT

The paper reviews certain integral equation approaches and related numerical methods used in studies of biomedical applications of electromagnetic fields pertaining to transcranial magnetic stimulation (TMS) and nerve fiber stimulation. TMS is analyzed by solving the set of coupled surface integral equations (SIEs), while the numerical solution of governing equations is carried out via Method of Moments (MoM) scheme. A myelinated nerve fiber, stimulated by a current source, is represented by a straight thin wire antenna. The model is based on the corresponding homogeneous Pocklington integro-differential equation solved by means of the Galerkin Bubnov Indirect Boundary Element Method (GB-IBEM). Some illustrative numerical results for the TMS induced fields and intracellular current distribution along the myelinated nerve fiber (active and passive), respectively, are presented in the paper.

DOI: 10.4018/978-1-7998-1204-3.ch013

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

While human exposure to radiation from different electromagnetic interference (EMI) sources has initiated a number of questions regarding potential adverse effects, particularly for the brain and eye exposure to high frequency (HF) radiation, some biomedical applications of electromagnetic fields are of particular importance, as well. Thus, the electromagnetic fields are applied in medical diagnostic and for therapy purposes featuring the use of techniques such as: transcranial magnetic stimulation (TMS) (Cvetković & et al., 2015; Garvey & Mall, 2008; Rajapakse & Kirton, 2013, Yamamoto & et al., 2016), percutaneous electrical nerve stimulation (PENS), transcutaneous nerve stimulation (TENS) (Frijnas & ten Kate, 1994; Rattay, 1999; Zulim & et al., 2015), or intraoperational methods such as transcranial electrical stimulation (TES) and direct cortical stimulation (DCS). Transcranial magnetic stimulation (TMS) is a noninvasive and painless technique for excitation or inhibition of brain regions, and in last few decades is an important tool in preoperative neurosurgical diagnostic/evaluation of patients (Picht & et al., 2013; Deletis & et al., 2014; Rogić & et al., 2014). TMS is also used in therapeutic purposes (i.e. depression), and is subject of interest in neurophysiologic research. Various efficiency aspects of TMS stimulation have been primarily stressed out primarily due to differences in relevant stimulation parameters such as pulse waveform, frequency, intensity of treatment, etc. The choice of optimal stimulation intensity is still investigated in many TMS studies.

The coil orientation and positioning appreciably influence the misalignment from the targeted brain region, thus reducing the TMS efficiency, although by using navigated TMS, this problem could be somewhat alleviated. In addition to the stimulation parameters, being adjustable to the requirements of the TMS operator, to a certain extent, the difference in individual brain morphology, due to age, gender or health status and the biological tissue parameters appreciably influence the distribution of the induced fields in the brain. Most of the parameters are obtained under different measurement on ex vivo animal and human tissues, and usually exhibit large variations from their average values. The level of uncertainty in the values of the brain conductivity and permittivity, is even more pronounced at low frequencies.

Modeling and computer simulation of TMS phenomena could be rather useful in determining the exact location of stimulation, in the interpretation of experimental results as well as in designing some more efficient stimulation setups. Realistic TMS models can also provide a more reliable prediction of the distribution of internal fields and currents by taking into account the variability of the various input parameters.

Furthermore, techniques such as PENS, electro-acupuncture, or TENS are widely used in a treatment of neurological disorders. Basically, there are two types of electric potential occurring in the stimulated nerve cells; the electrotonic potential and the action potential. The electrotonic potential existing due to the local changes in the ion conductivity decays along the fiber, and the passive membrane then shows linear nature which satisfies Ohm's law. The action potential is initiated when the threshold potential to which the membrane potential must be depolarized, is achieved. Studies on electrical excitation of nerves, among other aspects involve: nerve excitation using stimulating electrodes, nerve conduction velocity tests or non-invasive stimulation of nerves. The exceptional nervous system complexity, particularly nerve cells as its basis, and widespread neurological disorders and a need for better insight into complex functioning of the nervous system have been continuously motivating researchers (Golomolzina & et al., 2014) to carry out an efficient and accurate nerve fiber modeling. Thus, such computational models of nerve fibers could provide a study of the nerve fiber response to different stimulus waveforms, often used

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

within various electrotherapy techniques, particularly electro-acupuncture and PENS. Such a model can be beneficial and versatile tool for interpreting experimental results and analyzing variables for which there could be difficulties in the laboratory implementation.

The present work extends the review (Poljak & et al., 2016.a) of integral equation method approach to study TMS induced fields and nerve fiber excitation used in therapeutic procedures such as electro-acupuncture and/or PENS reported in (Poljak & et al. 2016.b) by providing more mathematical details and presenting some additional computational examples. The present work, together with (Poljak & et al. 2016.a) and (Poljak & et al. 2016.b) has been undertaken through tasks and activities within the framework of IEEE/ICES TC95 SC6 EMF Dosimetry Modeling and COST Action BM1309.

2. MODELING OF TMS INDUCED FIELDS

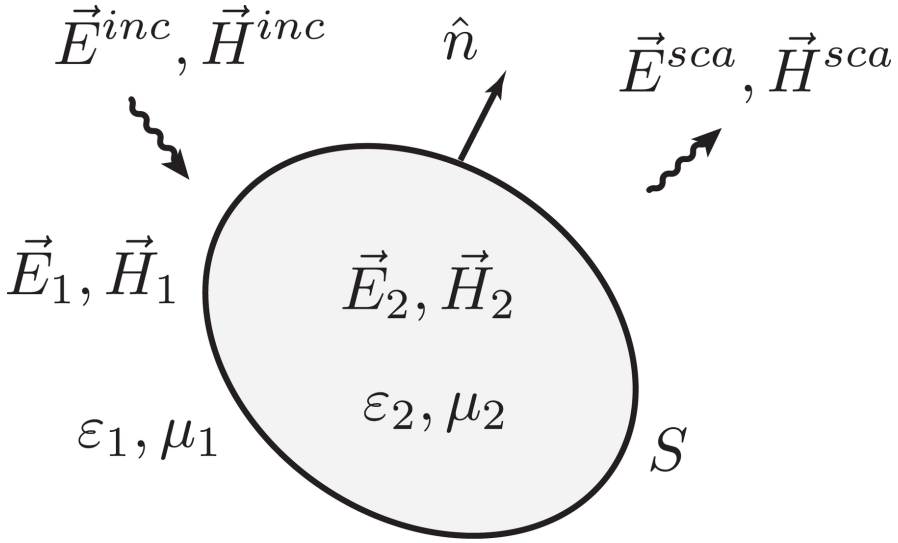
The lossy dielectric model of the brain exposed to TMS electromagnetic field is based on the set of coupled surface integral equations (SIEs) which could be derived from the equivalence theorem and by forcing the corresponding interface conditions for the electric and/or magnetic field, (Cvetković & et al., 2015) as indicated in Figure 1.

The lossy homogeneous object representing the brain is illuminated by the electromagnetic wave characterized by incident electric field $\vec{E}^{inc}$ and incident magnetic field $\vec{H}^{inc}$.

Performing some mathematical manipulations, the following set of coupled surface integral equations is obtained (Cvetković & et al., 2015):

$$\begin{aligned} &j\omega\mu_n \iint_S \vec{J}\left(\vec{r}'\right) G_n\left(\vec{r},\vec{r}'\right) dS' - \frac{j}{\omega e_n} \iint_S \nabla' s \cdot \vec{J}\left(\vec{r}'\right) \nabla G_n\left(\vec{r},\vec{r}'\right) dS' \\ &+ \iint_S \overrightarrow{M}\left(\vec{r}'\right) \times \nabla' G_n\left(\vec{r},\vec{r}'\right) dS' = \begin{cases} \vec{E}^{inc} & n = 1 \\ 0, & n = 2 \end{cases} \end{aligned} \tag{1}$$

Figure 1. The lossy homogeneous dielectric brain model



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

where $\vec{J}$ and $\vec{M}$ represent equivalent electric and magnetic current density, respectively, and G_n is the interior/exterior Green function given by (Cvetković & et al., 2015):

$$G_n\left(\vec{r},\vec{r}'\right)=\frac{e^{-jk_nR}}{4\pi R};\ R=\left|\vec{r}-\vec{r}'\right| \tag{2}$$

and R is the distance from the source to observation point, respectively, while k_n denotes the wave number of a considered medium n.

The brain model is generated from a freely available Google Sketchup rendering of the human brain, as shown on Figure 2.

The set of integral Equation (1) is solved by means of an efficient MoM scheme reported elsewhere, e.g. in (Cvetković & et al., 2015). For the sake of completeness, the MoM procedure is outlined in this section, as well.

As a first step, the equivalent electric and magnetic currents $\vec{J}$ and $\vec{M}$ in (1) are expressed by means of a linear combination of basis functions $\vec{f}_n$ and $\vec{g}_n$, respectively:

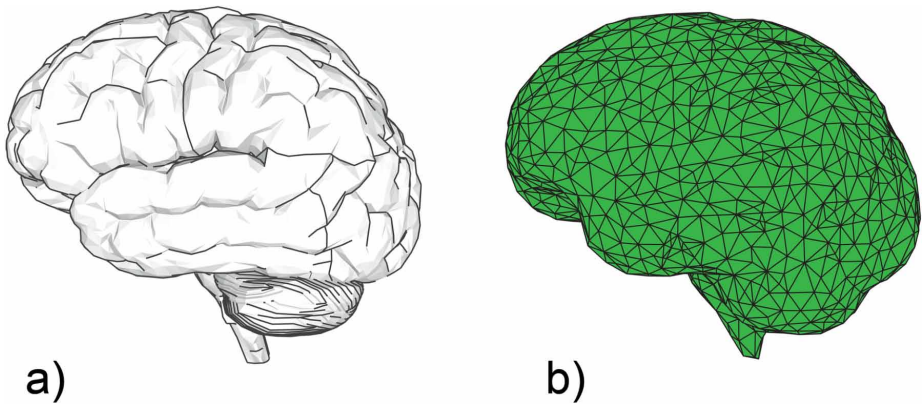
$$\vec{J}\left(\vec{r}\right)=\sum_{n=1}^{N}J_n\vec{f}_n\left(\vec{r}\right) \tag{3}$$

$$\vec{M}\left(\vec{r}\right)=\sum_{n=1}^{N}M_n\vec{g}_n\left(\vec{r}\right) \tag{4}$$

where J_n and M_n are unknown coefficients, while N is the total number of triangular elements.

Applying the weighted residual approach, i.e. multiplying (1) by the set of a test functions $\vec{f}_m$ and integrating over the surface S, after some mathematical manipulations, it follows:

Figure 2. The human brain model for SIE formulation: (a) Detailed 3-D model from Google Sketchup; (b) Final model discretized using the triangular elements



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$j\omega\mu_i\sum_{n=1}^{N}J_n\iint_S\vec{f}_m\left(\vec{r}\right)\cdot\iint_{S'}\vec{f}_n\left(\vec{r}'\right)G_i dS'dS+\frac{j}{\omega e_i}\sum_{n=1}^{N}J_n\iint_S\nabla s\cdot\vec{f}_m\left(\vec{r}\right)\iint_{S'}\nabla' s\cdot$$
$$\vec{f}_n\left(\vec{r}'\right)G_i dS'dS\pm\sum_{n=1}^{N}M_n\iint_S\vec{f}_m\left(\vec{r}\right)\cdot\left[\hat{n}\times\vec{g}_n\left(\vec{r}'\right)\right]dS+\sum_{n=1}^{N}M_n\iint_S\vec{f}_m\left(\vec{r}\right)\cdot\iint_{S'}\vec{g}_n\left(\vec{r}\right) \quad (5)$$
$$\times\nabla' G_i dS'dS=\begin{cases}\iint_S\vec{f}_m\left(\vec{r}\right)\cdot\vec{E}^{inc}, & i=1\\ 0, & i=2\end{cases}$$

where subscript *i* denotes the index of the medium. The details of the numerical solution procedure could be found elsewhere, e.g. in (Poljak & et al., 2016.a).

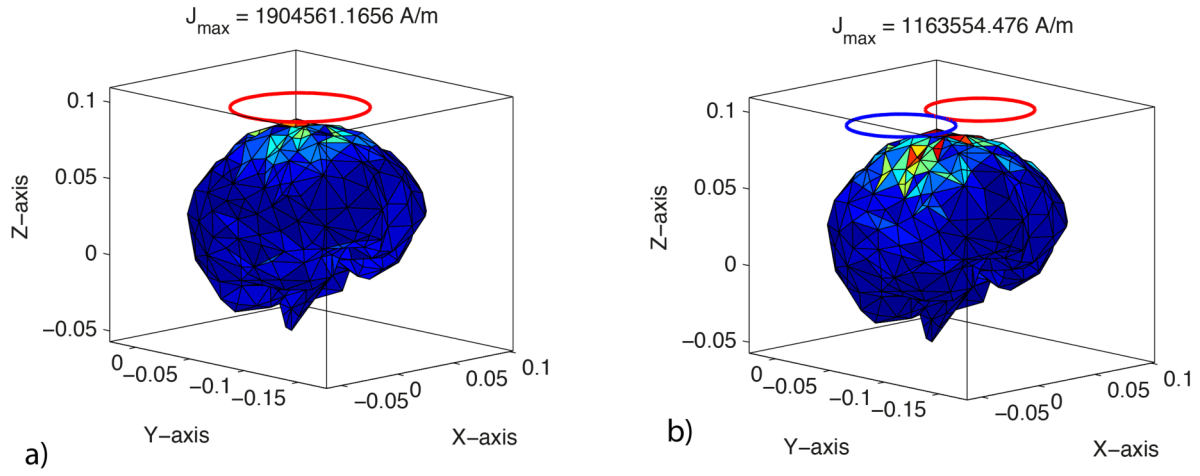
2.1. Numerical Results

Figures 3 and 4 show numerical results for the equivalent electric and magnetic current densities for the case of a) circular coil and b) figure-8 coil. The actual coil insulation and casing are not considered. Each coil is discretized into 80 linear segments. The coils are driven by a sinusoidal current of amplitude I = 2843 A, f=2.44 kHz. The radius of circular coil and figure-8 coil is 4.5 cm and 3.5 cm, respectively, while the number of windings for circular is 14 and for figure-8 is 15 turns.

Figures 5 and 6 show the numerical results for the sagittal and transversal cross-sections of the internal and external field, respectively, due to circular coil and figure-8 coil. In both cases, the coils are positioned tangentially to the brain surface, 1 cm over the primary motor cortex area, i.e. the distance between the coil geometric center and brain surface is 1 cm.

The results shown in Figures 5 and 6 confirm that the maximum electric field of figure-8 coil is induced directly under the coil geometric center, while for the case of circular coil, maximum field is induced under the coil windings. In addition to the induced electric field value, it is important to consider the orientation of the field vector. The results show the obvious importance of the correct TMS coil placement with respect to the brain surface.

Figure 3. Equivalent electric current density for: a) Circular coil; b) Figure-8 coil



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Equivalent magnetic current density for: a) Circular coil; b) Figure-8 coil

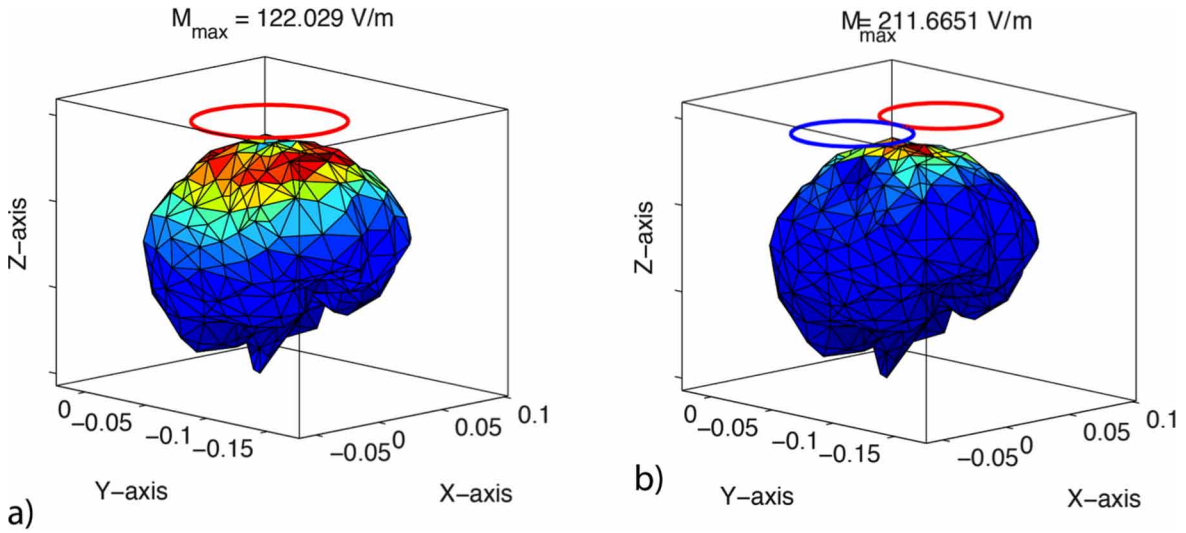
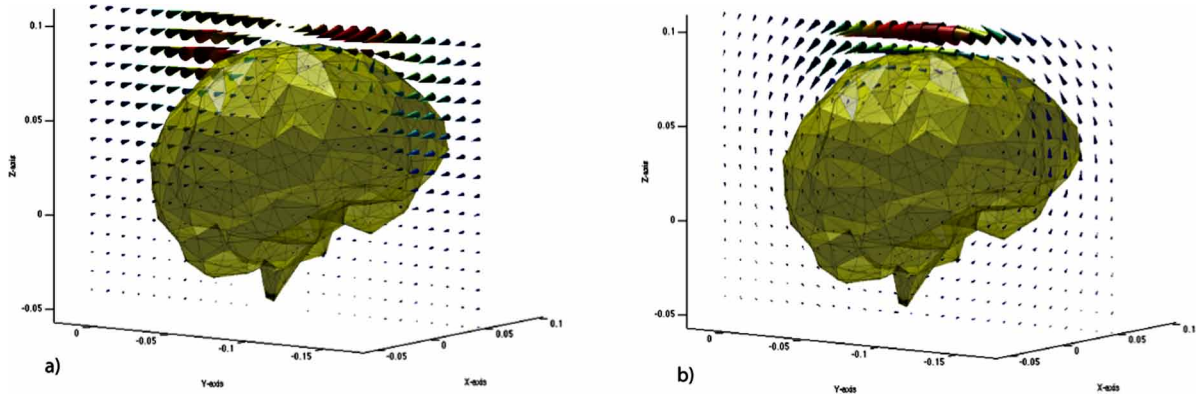


Figure 5. Sagittal cross-section of the electric field induced in the brain model by: a) Circular coil; b) Figure-8 coil. Coils not shown.



3. MODELING OF NERVE FIBER EXCITATION

Figure 7 shows the myelinated nerve fiber with an arbitrary number of Ranvier's nodes represented by a straight thin wire antenna (Zulim & et al., 2015; Zulim & et al. 2016).

Note that L denotes the fiber length, a is the inner axon radius and b is the radius including the myelin sheath. Electrode nerve fiber stimulation is taken into account by means of the equivalent current source I_g located at the fiber beginning. The current generator I_g represents the nerve fiber stimulation used in electro-acupuncture or percutaneous electrical nerve stimulation *(PENS)*, which both make use of thin needles injected through the skin. The fiber, oriented along the x axis, is assumed to be immersed in an unbounded homogenous lossy medium.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Transversal cross-section of the electric field induced in the brain model by: a) Circular coil; b) Figure-8 coil. Coils not shown.

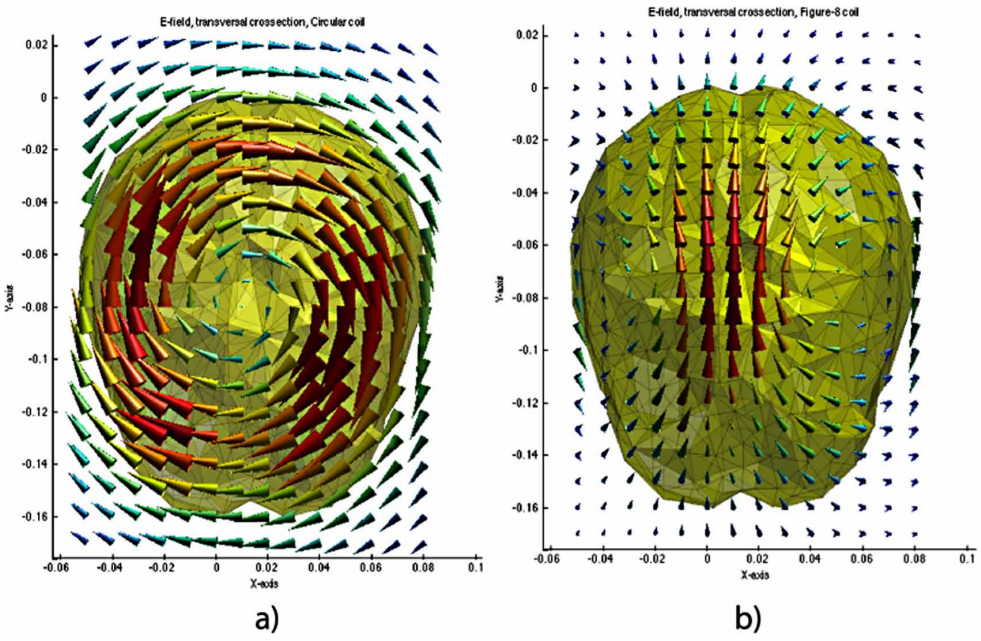
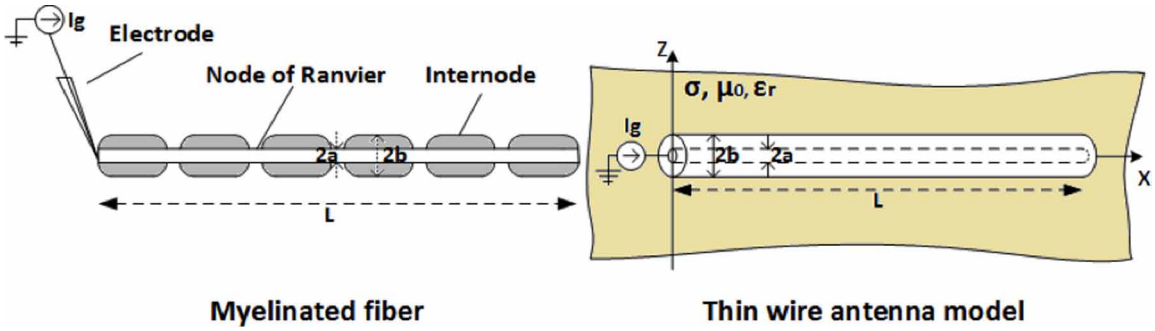


Figure 7. Thin wire antenna model of the myelinated nerve fiber



The formulation is based on the homogeneous Pocklington integro-differential equation for the unknown intracellular current I_a along the nerve fiber (Zulim & et al., 2015; Zulim & et al. 2016):

$$-\frac{1}{j4\pi\omega\varepsilon_{eff}}\int_0^L\left(\frac{\partial^2}{\partial x^2}-\gamma^2\right)g\left(x,x'\right)I_a(x')dx'=0 \tag{6}$$

while *g(x,x')* is the lossy medium Green's function defined as:

$$g\left(x,x'\right)=\frac{e^{-\gamma R}}{R} \tag{7}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

where R denotes a distance from the source to the observation point, respectively, γ is the complex propagation constant given by:

$$\gamma = \sqrt{j\omega\mu\sigma - \omega^2\mu\varepsilon_0\varepsilon_r} \tag{8}$$

and ε_{eff} is the complex permittivity of a lossy medium given by:

$$\varepsilon_{eff} = \varepsilon_0\varepsilon_r - j\frac{\sigma}{\omega} \tag{9}$$

with σ, ε_r and ω being the conductivity and relative permittivity of the medium, respectively, and ω is the angular frequency.

The current generator I_g is taken into account via the conditions at the nerve fiber ends:

$$I_a(0) = I_g,\ I_a(L) = 0 \tag{10}$$

It is worth mentioning that the properties of the lossy medium, nerve fiber membrane and myelin sheath are considered via the conductivity and relative permittivity which are obtained from the cable equation and the transmission line equation, respectively, as described in (Zulim & et al., 2015; Zulim & et al. 2016).

The Helmholtz equation for the corresponding transmembrane voltage is given by (Zulim & et al., 2015; Zulim & et al., 2016):

$$\frac{\partial^2 V}{\partial x^2} - \gamma_m^2 V = 0 \tag{11}$$

where:

$$\gamma_m = \sqrt{\frac{1 + j\omega\tau}{\lambda^2}} \tag{12}$$

and V is a transmembrane voltage, while λ is the length constant defined by (Zulim & et al. 2016):

$$\lambda = \sqrt{\frac{r_m a}{2\rho_a}} \tag{13}$$

and τ is the time constant given by (Picht & et al., 2013):

$$\tau = r_m c_m \tag{14}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ρ_a is the resistivity of axoplasm, r_m the myelin layer resistance for unit area and c_m capacitance of the Ranvier's node membrane per unit area.

The corresponding relations for the conductivity and relative permittivity are obtained by means of the procedure reported in (Zulim & et al., 2015). Homogeneous Pocklington integro-differential Equation (6) is solved via the Galerkin Bubnov Indirect Boundary Element Method (GB-IBEM). The details could be found elsewhere, e.g. in (Zulim & et al. 2016). A more general treatment of a thin-wire embedded in a lossy medium can be found in (Poljak & Doric, 2006.a, Poljak & Doric, 2006.b). For the sake of completeness, the procedure is outlined in this section, as well.

Thus, using the boundary element formalism, (6) is transformed into the set of linear equations, which can be written using the following matrix notation (Zulim & et al. 2016; Poljak, 2007):

$$\sum_{i=1}^{n}\left[Z\right]_{ji}^{e}\left\{\alpha\right\}_{i}^{e}=0,\ j=1,2,...n \tag{15}$$

where n is the number of wire segments, $\left[Z\right]_{ji}^{e}$ is the mutual impedance matrix representing the interaction of the observation segment j with the source segment i and $\left\{\alpha\right\}_{i}^{e}$ is the solution vector, expressed in global nodes. The mutual impedance matrix is:

$$\left[Z\right]_{ji}^{e}=-\frac{1}{j4\pi\omega\varepsilon_{eff}}\left[\int_{\Delta l_j}\int_{\Delta l_i}\left\{D\right\}_{j}\left\{D'\right\}_{i}^{T}g(x,x')-\gamma^{2}\int_{\Delta l_j}\int_{\Delta l_i}\left\{f\right\}_{j}\left\{f'\right\}_{i}^{T}g(x,x')dxdx'\right] \tag{16}$$

where matrices *{f}* and *{f'}* contain linear shape functions and matrices *{D}* and *{D'}* contain their derivatives. Δl_j and Δl_i are the observation and source segment lengths, respectively. The nerve fiber stimulation, in the form of an equivalent current generator, is incorporated into matrix system via boundary conditions (10).

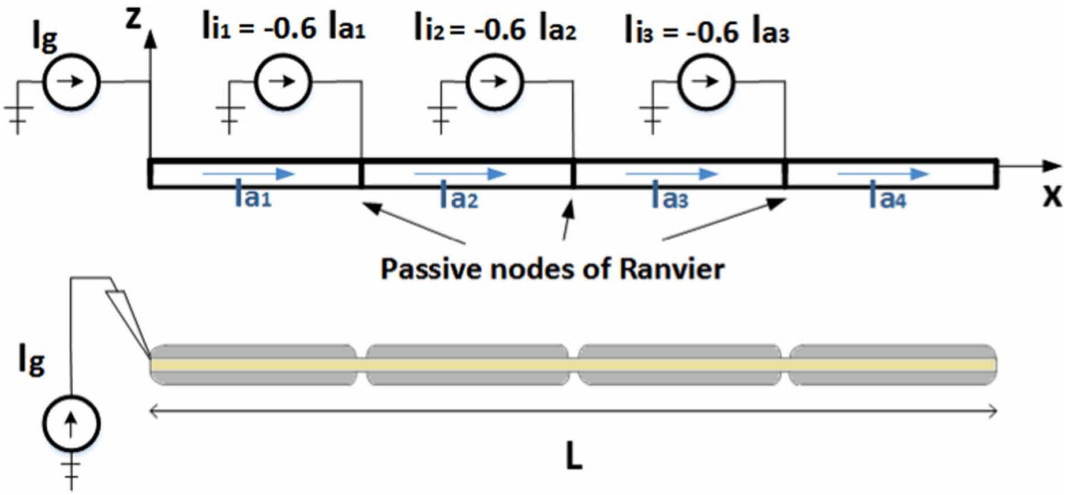
3.1. Passive Nerve Fiber

When modeling the passive nerve fiber, the ionic current in each Ranvier's node is to be considered. Thus, the additional current sources in the non-activated Ranvier's nodes are impressed. Figure 8 shows the passive nerve fiber model having 3 Ranvier's nodes and 4 internodes.

Analyzing the CRRSS (Chiu, Ritchie, Rogert, Stagg and Sweeney) model (Rattay, 1999; Chiu & et al., 1979; Sweeney & et al., 1987) it follows that the ionic current in each non-activated Ranvier's node is approximately equal to -0.6 I_a, where I_a is the intracellular current flowing into the observed Ranvier's node. Intracellular current I_{a2} which leaves the 1st Ranvier's node is equal to 0.4 I_{a1}. By analogy, the intracellular current I_{a3} flowing out of the second Ranvier's node is equal to 0.4 I_{a2} and the intracellular current I_{a4} that leaves the third Ranvier's node is *0.4* I_{a3}.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Passive nerve fiber model



3.2. Numerical Results for Passive Nerve Fiber

The intracellular current is calculated for the nerve fiber of 2 cm, with 9 Ranvier's nodes and 10 internodes. The fiber is stimulated at its beginning, using a subthreshold rectangular current pulse, plotted in Figure 9. The time domain results are obtained from the frequency domain results by means of Inverse Fast Fourier Transform (IFFT).

Ranvier's nodes are located at $x = 2, 4, 6, 8, 10, 12, 14, 16, 18$ mm along the fiber, respectively. The intracellular current distribution along the passive nerve fiber is plotted in Figure 10. The result is compared to the CRRSS model (Rattay, 1999; Chiu & et al., 1979; Sweeney & et al., 1987), and calculated for the time instant $t = 1$ ms.

Figures 11 and 12 show the subthreshold intracellular current versus time in the Ranvier's nodes 2, and 6, respectively.

The results for the intracellular current obtained by the antenna model, seem to be in a rather satisfactory agreement with the results obtained by using the CRRSS model. Note that the intracellular current waveform in Ranvier's nodes follows the excitation pulse form, as due to the subthreshold current stimulation, the voltage gated channels are not activated.

3.3. Active Nerve Fiber

Model of an active nerve fiber is presented on a myelinated nerve fiber with 3 active Ranvier's nodes and 4 internodes, as shown in Figure 13.

Each node is represented by a wire junction of two thin wires representing an active Ranvier's node. Thus, 3 additional current sources at the active Ranvier nodes represent an ionic current I_i of the activated node. The ionic current is determined by analyzing the *CRRSS* model and the analytical expression for the ionic current is obtained by curve fitting procedure (Poljak & et al., 2016.b):

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Rectangular subthreshold current pulse

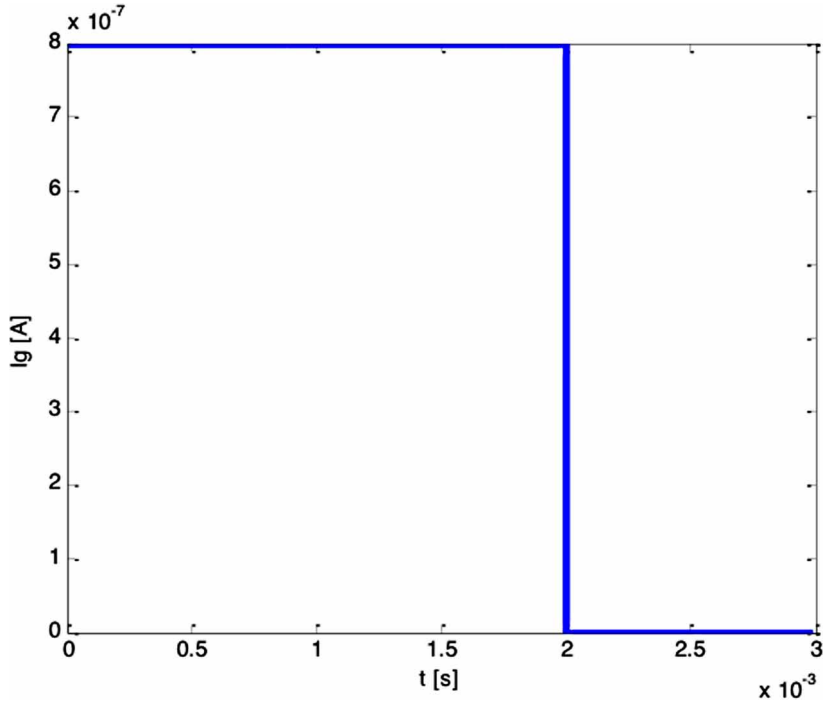
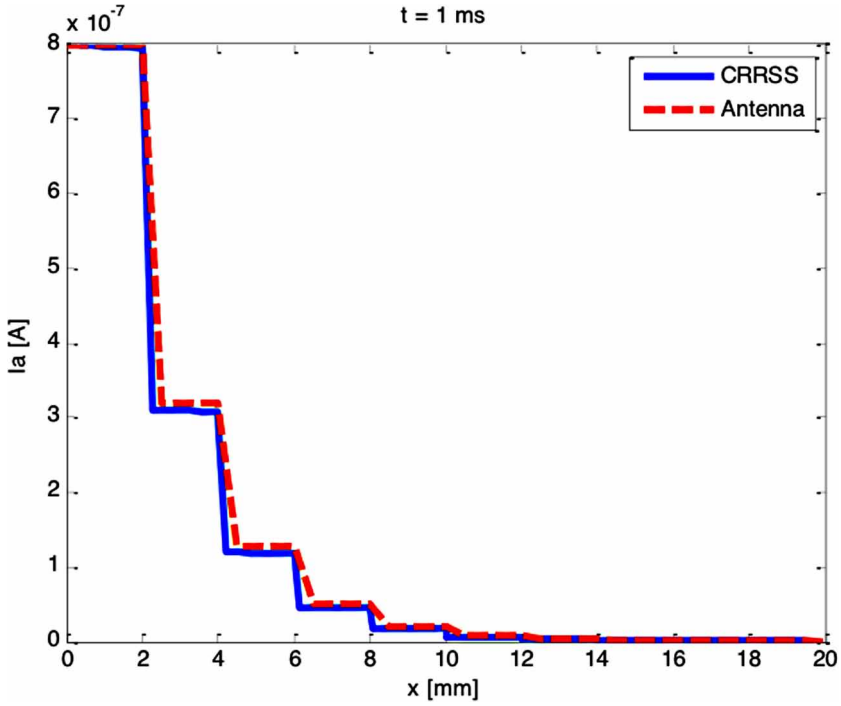


Figure 10. Intracellular current along the passive nerve fiber, t = 1ms, L=2cm



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Intracellular current in the passive Ranvier's node 2, L = 2 cm

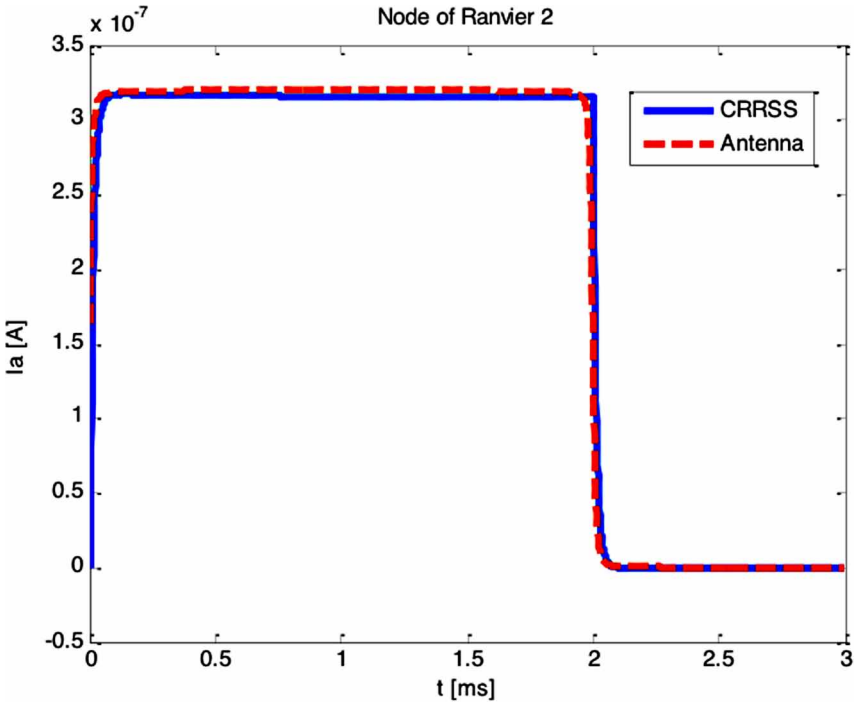
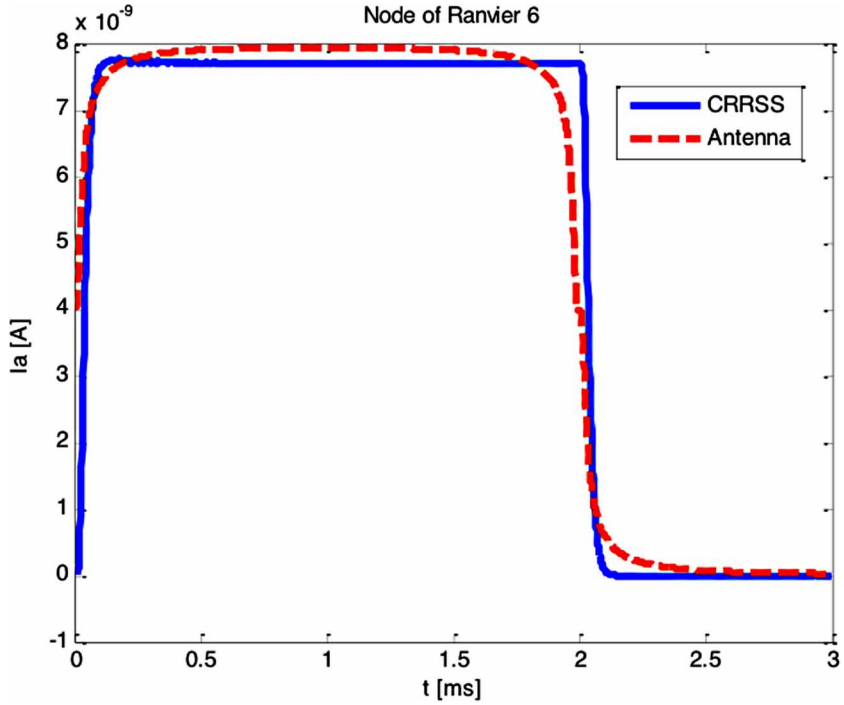
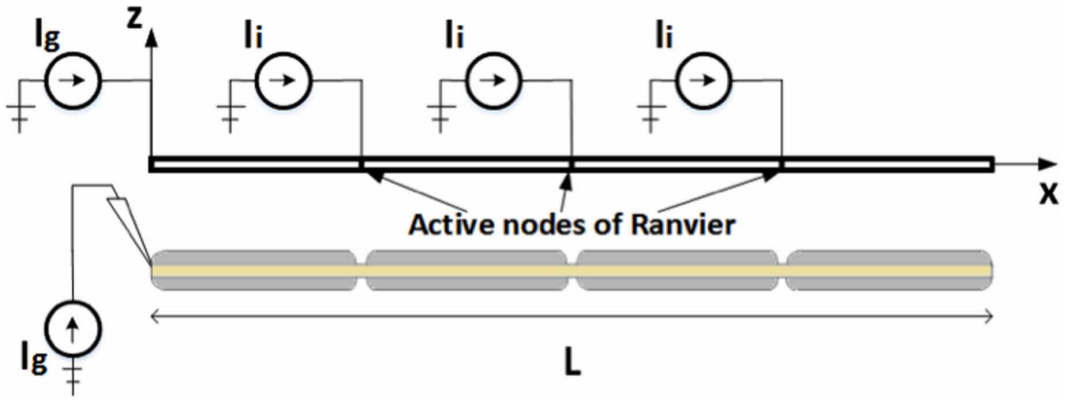


Figure 12. Intracellular current in the passive Ranvier's node 6, L = 2 cm



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 13. Active nerve fiber model



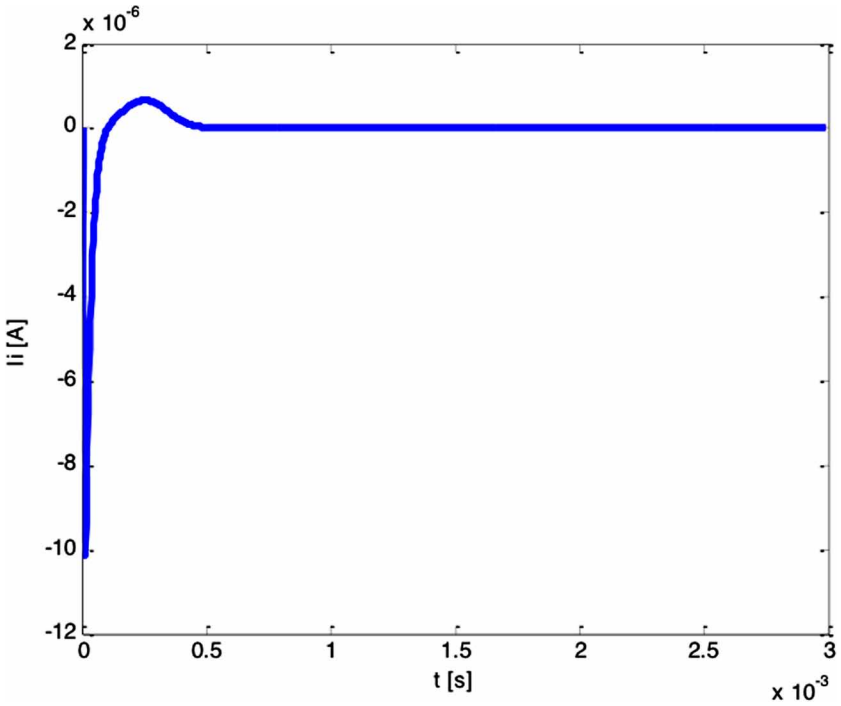
$$I_i(t) = -Au(t-t_1)\left[e^{-B(t-t_1)} - e^{-D(t-t_1)}\right] + Eu(t-t_1)e^{-G(t-t_2)^2} + He^{-K(t-t_3)^2} \tag{17}$$

where t_1 = 0.031 ms, t_2 = 0.252 msA, A = 0.02075 mA/ms, B = 199.9 (ms)$^{-1}$, D = 46.91 (ms)$^{-1}$, E = 0.000664 mA/ms, G = 60 (ms)$^{-2}$. The ionic current is depicted in Figure 14.

A thin wire junction model of the active node, can be represented by two separate thin wires, as shown in Figure 15.

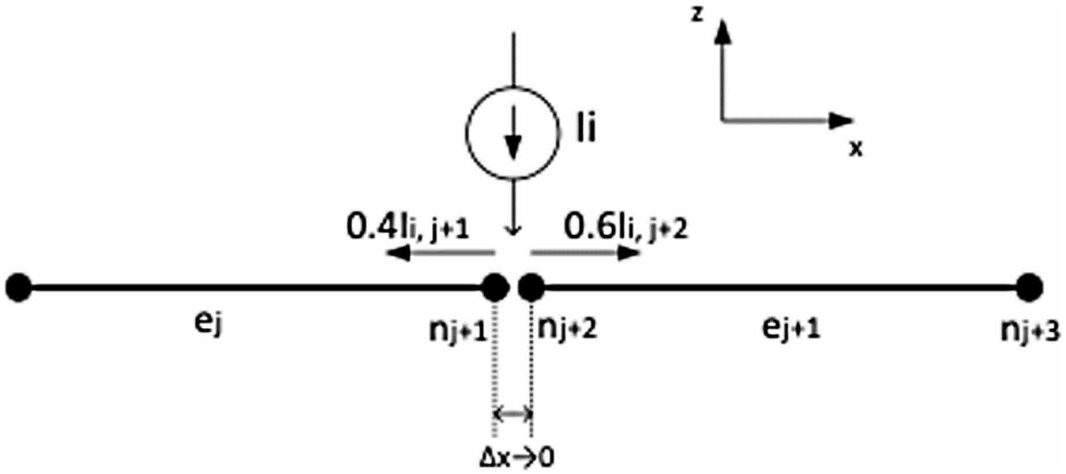
The Kirchhoff's law has to be satisfied at the junction (Poljak & et al., 2016.b):

Figure 14. Ionic current of activated node of Ranvier



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. Two wire junction representation of the active node



$$I_i = 0.4I_{i,j+1} + 0.61_{i,j+2} \tag{18}$$

where $I_{i,j+1}$ represents the ionic current value flowing out of the junction in one direction and $I_{i,j+2}$ is current flowing out of the junction in the opposite direction.

First, the nerve fiber is stimulated by a current generator I_g at a fiber beginning, as in the case of the passive nerve fiber. Furthermore, If the stimulating current exceeds the threshold in the corresponding Ranvier's node, the second current source, representing the ionic current, activates. The intracellular current for the activated fiber is a summation of the node activation current and the ionic current of the activated node. The resulting current, flowing out of the activated node, then activates the next node. The same procedure is repeated for each node.

The threshold for the nerve fiber activation depends on the strength and duration of the stimulus (Geddes & Bourland, 1985; Boinagrov & et al., 2010; Ashley & et al., 2005).

3.4. Numerical Results for Active Nerve Fiber

The intracellular current for the active nerve fiber is calculated with 9 Ranvier's nodes and 10 internodes. Rectangular current pulse used to depolarize the nerve fiber membrane is plotted in Figure 16.

The current pulse shown in Figure 16 stimulates the nerve fiber at the fiber beginning. Once a threshold current of 1.5μA reaches the first Ranvier's node, the other current source activates. Summation of the intracellular current activating the node and the ionic current of the second current source yields the resulting intracellular current in the 1st activated Ranvier's node. The resulting current then activates the 2nd Ranvier's node. The related intracellular current plot is shown in Figure 17.

The results for the intracellular current obtained via the antenna and CRRSS model show a rather satisfactory agreement to a certain extent. Some minor discrepancies appear due to the ionic current approximation. Intracellular current part pertaining to the rectangular current pulse is in a very good agreement.

Finally, Figure 18 shows the intracellular current for the 4th Ranvier's node.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. Rectangular superthreshold current pulse

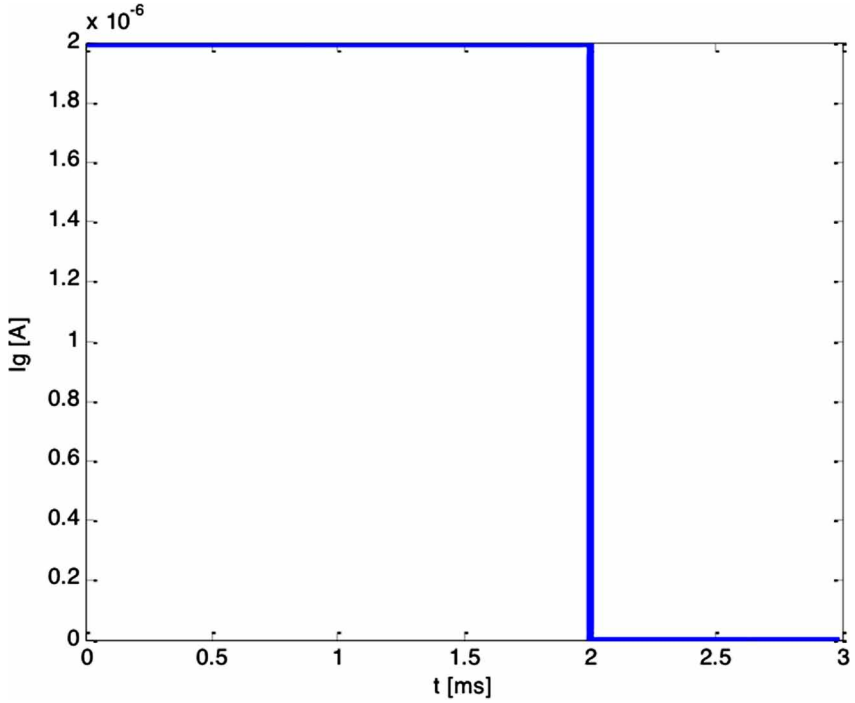
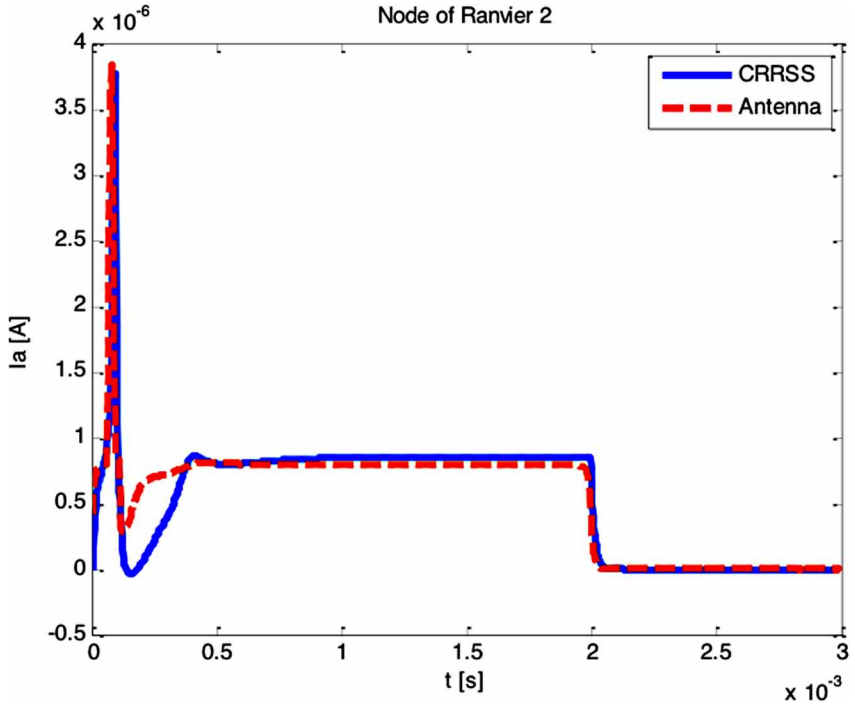


Figure 17. Intracellular current in the active Ranvier's node 2, L = 2 cm



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. Intracellular current in the active Ranvier's node 6, L = 2 cm

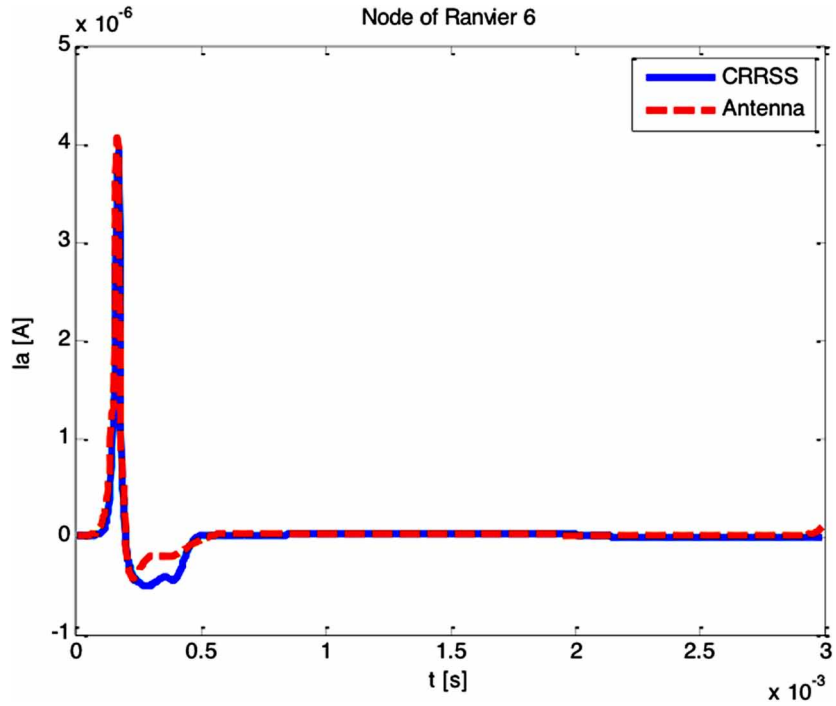
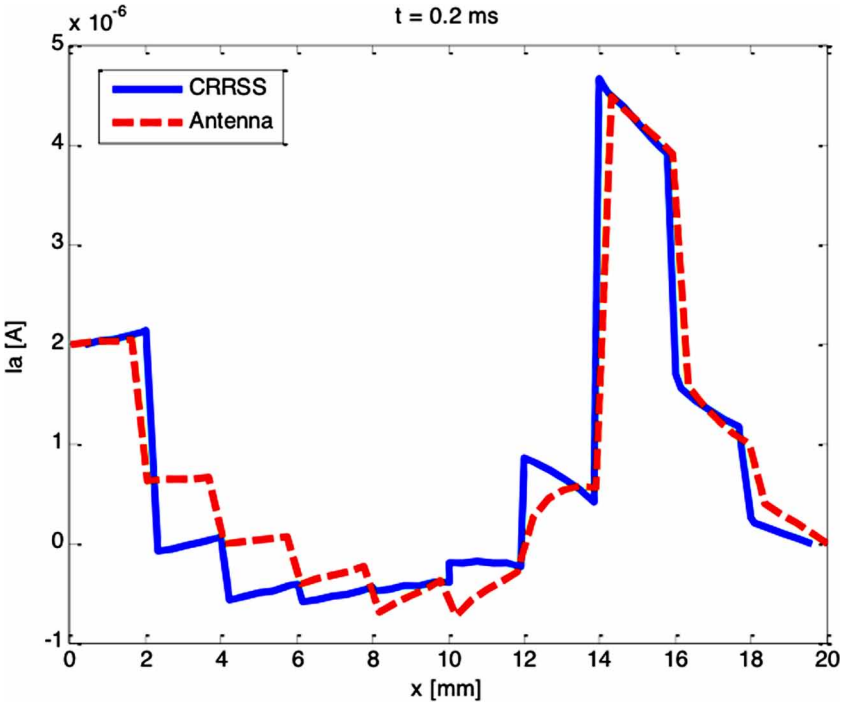


Figure 19. Intracellular current along the active nerve fiber, t = 0.2 ms, L = 2cm



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Obviously, the rectangular current pulse decays along the fiber, while intracellular current part keeps almost the same amplitude at the activated node. The signal of the activated nerve propagates along the fiber almost without attenuation.

Finally, Figure 19 shows the intracellular current distribution along the nerve fiber, plotted for the time instant of 0.2 ms obtained by the antenna and CRRSS model, respectively.

There is a satisfactory agreement between the results, with some visible discrepancies in amplitude of the intracellular current travelling along the fiber, mainly due to the instability of numerical procedure in the antenna model.

4. CONCLUSION

The paper reviews the use of integral equation methods in studies of biomedical applications of electromagnetic fields pertaining to TMS and nerve fiber stimulation. TMS modeling is based on the set of coupled surface integral equations while related numerical solution procedure based on the Method of Moments (MoM) approach.

The integral equation based formulation presents a more rigorous, and hence more accurate description of the TMS, by not having to take into account the inductive and capacitive effects, often being neglected when quasi-static approximation is used.

The myelinated nerve fiber is modeled by the thin wire antenna in passive and active state, respectively, stimulated by a rectangular current pulse at the fiber beginning. The model is based on the corresponding Pocklington integro-differential equation, handled via the GB-IBEM.

The myelinated nerve fiber based on antenna theory presents the completely new concept in the nerve fiber modeling. Interested reader can find the detailed survey of other concepts related to numerical electrostimulation models in (Reilly 2016).

Some illustrative computational examples for the induced field in the brain, due to the TMS treatment and intracellular current along the nerve fiber due to the current source excitation are presented in the paper.

The future work will be related to the development of anatomically realistic model of the human brain based on the integral equation approach. Also, the coupling of the human brain model with the nerve fiber model, i.e. the induced electric field results as the input parameters of the nerve fiber model is the subject of the ongoing activities.

REFERENCES

Ashley, Z., Sutherland, H., Lanmuller, H., Unger, E., Li, F., Mayr, F., ... Salmons, S. (2005). Determination of the chronaxie and rheobase of denervated limb muscles in conscious rabbits. *Artificial Organs*, *29*(3), 212–215. doi:10.1111/j.1525-1594.2005.29037.x PMID:15725219

Boinagrov, D., Loudin, J., & Palanker, D. (2010). Strength–Duration Relationship for Extracellular Neural Stimulation: Numerical and Analytical Models. *Journal of Neurophysiology*, *104*(4), 2236–2248. doi:10.1152/jn.00343.2010 PMID:20702740

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chiu, S. Y., Ritchie, J. M., Rogart, R. B., & Stagg, D. (1979). A quantitative description of membrane currents in rabbit myelinated nerve. *The Journal of Physiology*, *292*(1), 149–166. doi:10.1113/jphysiol.1979.sp012843 PMID:314974

Cvetković, M., Poljak, D., & Haueisen, J. (2015). Analysis of Transcranial Magnetic Stimulation Based on the Surface Integral Equation Formulation. *IEEE Transactions on Biomedical Engineering*, *62*(6), 1535–1545. doi:10.1109/TBME.2015.2393557 PMID:25608302

Deletis, V., Rogić, M., Fernández-Conejero, I., Gabarrós, A., & Jerončić, A. (2014). Neurophysiologic markers in laryngeal muscles indicate functional anatomy of laryngeal primary motor cortex and premotor cortex in the caudal opercular part of inferior frontal gyrus. *Clinical Neurophysiology*, *125*(9), 1912–1922. doi:10.1016/j.clinph.2014.01.023 PMID:24613682

Frijns, J. H., & ten Kate, J. H. (1994). A model of myelinated nerve fibres for electrical prosthesis design. *Medical & Biological Engineering & Computing*, *32*(4), 391–398. doi:10.1007/BF02524690 PMID:7967803

Garvey, M. A., & Mall, V. (2008). Transcranial Magnetic Stimulation in Children. *Clinical Neurophysiology*, *119*(5), 973–984. doi:10.1016/j.clinph.2007.11.048 PMID:18221913

Geddes, L. A., & Bourland, J. D. (1985). The Strength-Duration Curve. *IEEE Transactions on Biomedical Engineering*, *BME-32*(6), 458–459. doi:10.1109/TBME.1985.325456 PMID:4007910

Golomolzina, D. R., Gorodnichev, M. A., Levin, E. A., Savostyanov, A. N., Yablokova, E. P., Tsai, A. C., ... Smirnov, N. V. (2014). Advanced Electroencephalogram Processing: Automatic Clustering of EEG Components. *International Journal of E-Health and Medical Communications*, *5*(2), 49–69. doi:10.4018/ijehmc.2014040103

Picht, T., Krieg, S. M., Sollmann, N., Rösler, J., Niraula, B., Neuvonen, T., ... Ringel, F. (2013). A comparison of language mapping by preoperative navigated transcranial magnetic stimulation and direct cortical stimulation during awake surgery. *Neurosurgery*, *72*(5), 808–819. doi:10.1227/NEU.0b013e3182889e01 PMID:23385773

Poljak, D. (2007). *Advanced Modeling in Computational Electromagnetic Compatibility*. New Jersey: Wiley-Interscience. doi:10.1002/0470116889

Poljak, D., Cvetković, M., Dorić, V., Zulim, I., Đogaš, Z., Rogić-Vidaković, M., . . . El Khamlichi Drissi, K. (2016b). Integral Equation Models in Some Biomedical Applications of Electromagnetic Fields: Transcranial magnetic stimulation (TMS), Nerve fiber stimulation. In *Proceedings of the International Multidisciplinary Conference on Computer and Energy Science (SpliTech),* Croatia. 10.1109/SpliTech.2016.7555936

Poljak, D., Cvetković, M., Peratta, A., Peratta, C., Dodig, H., & Hirata, A. (2016a, June). On some integral equation approaches in electromagnetic dosimetry. In Proceedings of BioEM 2016, Ghent Belgium.

Poljak, D., & Doric, V. (2006). a). Wire antenna model for transient analysis of simple grounding systems, Part I: The vertical grounding electrode. *Progress in Electromagnetics Research*, *64*, 149–166. doi:10.2528/PIER06062101

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Poljak, D., & Doric, V. (2006). b). Wire antenna model for transient analysis of simple grounding systems, Part II: The horizontal grounding electrode. *Progress in Electromagnetics Research*, *64*, 167–189. doi:10.2528/PIER06062102

Rajapakse, T., & Kirton, A. (2013). Non-invasive brain stimulation in children: Applications and future directions. *Transl. Neurosci.*, *4*(2), 1–29. doi:10.247813380-013-0116-3 PMID:24163755

Rattay, F. (1999). The basic mechanism for the electrical stimulation of the nervous system. *Neuroscience*, *89*(2), 335–346. doi:10.1016/S0306-4522(98)00330-3 PMID:10077317

Rogić, M., Deletis, V., & Fernández-Conejero, I. (2014). Inducing transient language disruptions by mapping of Brocas area with modified patterned repetitive transcranial magnetic stimulation protocol. *Journal of Neurosurgery*, *120*(5), 1033–1041. doi:10.3171/2013.11.JNS13952 PMID:24405070

Sweeney, J. D., Mortimer, J. T., & Durand, D. (1987). Modeling of mammalian myelinated nerve for functional neuromuscular electrostimulation. In *Proc 9th Ann Conf IEEE EMBS* (pp. 1577–1578).

Yamamoto, K., Takiyama, Y., Saitoh, Y., & Sekino, M. (2016). Numerical analyses of transcranial magnetic stimulation based on individual brain models by using a scalar-potential finite-difference method. *IEEE Transactions on Magnetics*, *52*(7), 1–4.

Zulim, I., Dorić, V., & Poljak, D. (2016). Myelinated nerve fiber antenna model activation. In *Proceedings of the 2016 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE. 10.1109/SOFTCOM.2016.7772159

Zulim, I., Dorić, V., Poljak, D., & El Khamlichi Drissi, K. (2015). Antenna model for passive myelinated nerve fiber. In *Proceedings of the 2015 23rd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE. 10.1109/SOFTCOM.2015.7314069

This research was previously published in the International Journal of E-Health and Medical Communications (IJEHMC), 9(1); edited by Joel J.P.C. Rodrigues; pages 65-84; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 14
Efficient Algorithm for Answering Fuzzy Medical Requests in Pervasive Healthcare Information Systems

Wissem Labbadi
University of Tunis, Tunisia

Jalel Akaichi
King Khalid University, Saudi Arabia

ABSTRACT

The progress in mobile devices and wireless networks technologies has considerably contributed to integrate pervasive computing expertise in many domains with the aim of improving the quality of services and users' mobility. However, in many situations, users may face difficult situations, needing faster decisions, where classical systems impose the submission of classic queries in which crisp conditions must be carefully fixed. This inconvenience limits the potential of pervasive applications accessed by users having few times to make the right decisions. To introduce the contributions of this paper, we choose the medical domain as example. We considered a pervasive healthcare application under which physicians haven't enough time to fix carefully their queries in some emergency cases. Therefore, they are allowed to flexibly express their preferences using conjunctive fuzzy queries and to quickly receive best answers anywhere and anytime while treating patients in the shortest time and consequently free resources for eventually other urgent requests. In this work, we consider, in general, the problem of efficiently finding the top-K answers for a conjunctive fuzzy query from the top-N conjunctive query rewritings of the query. In particular, we propose an efficient algorithm called the Top-N rewritings algorithm for finding the top-N query rewritings of a medical conjunctive fuzzy query using a set of conjunctive crisp views. At the best of our knowledge, this algorithm is the first to generate, without computing all possible rewritings, the N best ones ordered according to their satisfaction degrees and that are likely to return the best K-answers for the user fuzzy query. The relevance of a query rewriting is estimated using a second algorithm called the Query-satisfaction computing algorithm proposed to estimate, through the histograms maintained to approximate the distribution of set of values returned by the rewriting and to which fuzzy predicates are related, the pertinence of a conjunctive fuzzy query rewriting rather than accessing the database relations.

DOI: 10.4018/978-1-7998-1204-3.ch014

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The integration of pervasive computing in healthcare has enabled patients to take profit from high quality health care services anytime and anywhere. For this reason, providing continuous remote monitoring especially for elderly patients into the comfort of their home will be certainly less expensive than the cost of in-clinic recovery and more efficient. Furthermore, the advancement in mobile devices and in context-aware technologies has contributed to enhance the quality of healthcare services and the users' mobility from one place to another.

Pervasive healthcare has the potential to reduce medical costs and improve quality of services; however it faces many technical difficulties related to the integration of the different medical information systems interoperating under the same healthcare application and maintaining periodically changeable schemas. In front of this challenging problem, middleware is gaining interest by insulating different systems from proprietary interfaces and providing layers of abstraction giving hence the illusion of using a homogeneous system. In many cases, physicians face problematic situations threatening patients' lives and needing fast and firm decisions; nonetheless classical medical middleware systems used to force physicians to submit traditional queries in which they must inflexibly precise their crisp conditions. This inconvenience limits the potential of the healthcare application used by physicians having few times to save patients' lives. In fact, they might have enough time to fix carefully their queries and can't wait any longer for needed answers. For this reason, we put forward a pervasive healthcare application based on an efficient virtual integration system in order to allow physicians to express freely their preferences using natural language and vague terms and consequently quickly to obtain their desirable answers anywhere and anytime while treating patients in the shortest time and therefore free medical resources for eventually other urgent requests.

In this paper, we propose the Top-N rewritings algorithm, a novel and efficient algorithm for generating ordered top-N query rewritings which are likely to return the top k-tuples fitting with user preferences. This algorithm is the first one to generate the N best rewritings without computing all possible ones while it discards useless rewritings from the query rewritings searching step and hence, generates only the best likely satisfactory ones ordered according to their satisfaction degrees of the user query. We discuss also two different solutions on how exploiting the output of the Top-N rewritings algorithm in order to recover the Top-K tuples considered as best answers. Finally, we envision an implicit query cache memory in the middleware to avoid useless query processing when it is already available in the cache.

In the next section, a brief summary about pervasive healthcare is presented. Section 3 shows the contribution of middleware to help integrating different medical information systems interoperating under the same healthcare application. In section 4, we define a fuzzy medical conjunctive query. Section 5 is devoted to describe the two proposed algorithms. This section describes also how to exploit the Top-N rewritings to determine the Top-K tuples. Section 6 explains when and how using the query memory cache. In section 7 we discuss the feasibility of the study andthe final section concludes the paper.

Pervasive Healthcare Applications

The progress in mobile devices over wireless networks and in context-aware technologies has considerably contributed to integrate pervasive computing in healthcare domain (Bardram, 2004; Bardram & Christensen & Olsen, 2004; Christensen & Bardram, 2002; Favela & Rodriguez & Preciado & Gonzalez, 2004; Varshney, 2003). In fact, by using wireless mobile devices, physicians become able to better com-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

municate with patients, access updated patient medical records anywhere and anytime, and furthermore to share and exchange medical information with other colleagues in order to provide patients with high quality health care services into medical institutions and/or at the comfort of their own homes.

Pervasive healthcare focuses on the use of pervasive computing technology for developing services to improve quality of medical care. Carrying out a remote consultation or providing continuous remote monitoring especially for elderly patients in their post-operative period into the comfort of their home will be less expensive than the cost of in-clinic recovery and patients are becoming independent from stationary treatment. In addition, pervasive computing aims to higher the quality of healthcare services although user's mobility, so to cope with changing context due to mobility of the user in the environment, applications need to be context-aware in order to identify their locations and assuring real time access to the mobile devices to bring them information about the nearest hospitals, the nearest pharmacies, the nearest emergencies while moving from one region to another. Furthermore, it helps optimize physician meetings. For example, if a doctor needs an in-home help of his colleagues, it is better to choose them between the nearest ones. A wide range of pervasive healthcare systems and smart devices have been proposed over the few past years. In (Shahriyar, Bari, Kundu, Ahamed, & Akbar, 2000), authors present a bio-sensor based mobile health monitoring system named ″Intelligent mobile health monitoring system (IMHMS)″ which improve health monitoring by providing medical feedback to the patients anytime anywhere through mobile devices based on the medical information collected intelligent and miniature bio-sensor nodes, that can be easily integrated into wireless personal or body area networks for mobile health monitoring, which are capable of sensing, processing, and communicating. Many health care projects are developed, before the appearance of the IMHMS, with the objective of providing more and more assistance to the elderly. For example, the project named AHRI (Aware Home Research Initiative) (Ahamed, Haque, Stamm, & Khan, 2007) with the goal is to provide a unified solution for the seniors in the home, enabling them to closely participate in disease detection and health management by themselves. (Halteren et al., 2004) Developed a system named Wellness Assistant (WA) which can be used, through inexpensive handheld devices, by people with obesity, diabetes, or high blood pressure, conditions which need constant monitoring. In (Varshney, 2007), Varshney describes various pervasive healthcare applications and some issues and challenges to the use of wireless technology in the task of monitoring. Varshney is considered among the initiates to work on the integration of wireless networks technologies in healthcare that he addresses and discusses how wireless technologies can be applied in the healthcare environment and how to effectively manage the complexity of wireless and mobile networks in healthcare applications. In (Varshney, 2006a), he discusses and illustrates how mobile and wireless technologies can be applied for location management, intelligent emergency system, patient monitoring and mobile telemedicine applications. He also designed some wireless infrastructure for some future healthcare applications. In (Varshney, 2006b), Varshney determines several of the requirements of using wireless networks and mobile technologies in patient monitoring and show how infrastructure oriented wireless LANs can be used to support patient monitoring in diverse environments.

Some other related work propose to integrate pervasive healthcare into homes in order to be smart homes that allow patients to have a better quality of life due to fact that is a familiar environment. In (Bartolomeu, Fonseca, & Vasques, 2008), authors explain the requirements needed to build a smart home to support the monitoring and care for chronically ill or recovery patients. They describe two systems of Pervasive healthcare in smart homes. Altcare (Sohaib, Elbrandt, Dragon, & Ostermann, 2010) is able to detect the daily events of an elderly using a network of cameras. This system detects emergency situations and automatically contacts the responsible for the patient and also allows doctors at any time to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

check the patient's condition and thus provides a more comfortable and safe living to the elderly. Smart Condo (Boers, et al., 2009) is a monitoring system for elderly at home with the characteristic that it respects the privacy of individuals by not recording any videos or sound.

An other important work which uses context-aware devices in healthcare is the AWARE project (Bardram, Hansen, Mogensen, & Sogaard, 2006), which implements a context-aware environment in a hospital. The system is essentially composed of a tracking module for clinics and patients, an infrastructure for acquisition, management and distribution of contextual information, and tools like situated displays and mobile phones.

An important recent work which illustrates the interoperability of heterogeneous IoT (Internet of Things) platforms is (Gaglio, Re, Martorella, Peri, & Vassallo, 2014) in which authors developed a mobile health system that enables to monitor distant patients. Such a system based on high-level programming processing can address some important issues arising when heterogeneous devices are requested to cooperate, which is a crucial feature of IoT applications (Lui, Antonio, & Giacomo, 2010).

Finally, (Nitin & Sujata, 2015) proposes a pervasive monitoring of mobile health care using Android for mobile healthcare emergency. With pervasive monitoring, smart phone resources including computing power and energy can be opportunistically gathered to process the computing intensive personal health information during mobile healthcare emergency with minimal privacy disclosure.

Pervasive healthcare has the prospective to reduce medical costs and to increase the quality of services, but then again it likewise faces many technical hitches related to the integration of the diverse medical information systems interoperating under the same healthcare applications. In fact, most of existing implementations of pervasive healthcare applications do not interoperate well due to segmented heterogeneous and autonomous solutions (Black & Segmuller & Cohen, 1995). The different information systems treat various types of information that evolves dynamically which leads to the changing of the underlying schemas of the database and the interfaces dedicated to the other integrated organization systems. The problem persists in how the back-end system managing the different integrated information systems might evolve dynamically?

Middleware Pervasive Healthcare Applications

To face the problem of integrating autonomous medical information systems handling different types of medical information that evolves dynamically incurring consequently the change in the interface offered to the other integrated medical information systems, middleware was gaining interest by insulating different systems from proprietary interfaces and providing layers of abstraction giving hence the illusion of using a homogeneous system. Many prototypes have been developed to ease the design and the integration of the different information systems behind the fine running of the pervasive healthcare applications. In this context, building a pervasive healthcare application based on a virtual information integration system (Lenzerini, 2002) was very promising to integrate medical databases. The system follows the local as view approach where the distributed medical information sources are defined as views over a global schema and a middleware receiving and answering users' queries. Such approach favors the addition of new medical data sources such as new hospitals, pharmacies, emergencies, etc. A virtual medical information integration system provides the healthcare application with a unified schema giving the illusion to the application users of using a unified system without caring about the heterogeneity between the different information sources. For example, a physician may need the recent analysis results and medicines taken by a patient. In this case, the required information must be extracted from

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

different hospitals and pharmacies databases and joined together. That is the role of mediation system which deals with the problem of gathering information from the different medical information sources by finding correspondences between the global and the data sources schemas.

In many cases, physicians face difficult situations related to patients' states needing fast decisions; however classical medical integration systems usually force them to submit classic queries in which they must express precisely their crisp conditions. This inconvenience, obviously, limits the help expected from the healthcare application used by physicians having few time to react according to the patients' states. In fact, they, usually, haven't enough time to fix carefully their queries and couldn't wait any longer for answers.

Fuzzy Medical Conjunctive Query

A fuzzy medical conjunctive query is a conjunctive query (Ullman, 2000) in which fuzzy predicates are related to medical variables such as blood pressure, heartbeat, blood glucose level, etc. A conjunctive query is defined as a rule applied to the extensional database or stored relation by considering all possible substitutions of values for the variables in the body of the query. If a substitution makes all the sub goals true, then the same substitution, applied to the head of the query, is an inferred fact about the head's predicate. A fuzzy medical conjunctive query Q is supposed to respect the following form:

$$Q(X):- \exists\, Y \text{ body } (X, Y), p_1(x_1), \ldots, p_n\,(x_n) \quad (1)$$

Where X are the distinguished variables, Y are so-called the non-distinguished variables and are distinct from the variables in X, body(X, Y) is a conjunction of non-fuzzy predicates, p_i is a fuzzy predicate, related to a single distinguished medical variable (x_i), which assigns to each tuple t_i a score $\mu p_i\,(t_i) \in [0, 1]$ describing the satisfaction degree of the fuzzy predicate p_i by the tuple t_i according to a given membership function.

Let's consider the following relations:

- Patient (PatID, name, age)
- Blood_characteristics (PatID, blood_pressure, blood_glucose_level, time)

An example of a conjunctive fuzzy query is ''Show names of young patients where blood pressure is high''. This query can be formulated, using logical rule notation, as follows:

Answer (y): Patient (x, y, z), Blood_characteristics (x, v, w, a), young (z), high (v)

Flexible and Fast Pervasive Healthcare Applications

To remedy these inconveniences, more flexibility in the definition of user queries is required. In fact, to get the best matched answers, medical applications must allow physicians to express freely their preferences using natural language enhanced by vague or imprecise conditions in their queries, especially when the limits of their conditions are not clear. Therefore, to help users submitting their requests, many fuzzy conditions over many attributes are envisioned, and sers have only to choose the appropriate desired conditions. Moreover, if a user requirement is not defined yet, it can be added, anytime, to the requirement list for an eventual future use. In addition, it is possible to use fuzzy quantifiers inside

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

queries like "many" and "at least" in fuzzy quantified statements. Basing on flexibility advantages of the fuzzy queries, users will no longer care about the exact limits of their conditions in urgent cases when such details can be forgotten.

A simple way to answer a conjunctive fuzzy query submitted to the medical information integration system is to apply Bucket (Levy & Rajaraman & Ordille, 1996), MiniCon (Pottinger & Halevy, 2000) or Inverse rules (Dushka & Genesereh & Levy, 2000) algorithms for answering classical queries using views, in order to determine all possible rewritings which can be sorted according to their computed satisfaction degrees (Labbadi & Akaichi, 2010) and materialize only those with accepted scores. However, this method incurs an unnecessary cost since it determines also the useless query rewritings and causes a waste of resources for useless scores computation. To overcome these drawbacks, we propose a method based on two steps. In the first step, we propose the *Top-N rewritings algorithm* which generates the N best likely satisfactory rewritings ordered according to their satisfaction degrees of the user query. The relevance of a query rewriting is estimated using the *Query-satisfaction computing algorithm* proposed to estimate the pertinence of a fuzzy conjunctive query rewriting rather than accessing the database relations. This algorithm, proposed in (Labbadi & Akaichi, 2010), estimate the satisfaction degree of a fuzzy predicate by a tuple from a materialized view defined as a simple fuzzy selection query of the form "X is P" where P is a fuzzy condition related to attribute X. The estimations are obtained from the histogram constructed dynamically to approximate the frequency distribution of X in the resulting relation.

In the second step, the output of the Top-N rewritings algorithm is used to find the Top-K tuples for the submitted fuzzy query. We consider our solution as an efficient one able to answer medical fuzzy queries because it intends to accelerate the medical integration system processing. In fact, in the first step, only the N best rewritings are determined without computing or searching for all them to evaluate their interest and in the second step, only the best rewritings are used to find the Top-K tuples for users' queries.

In the following, we propose the query satisfaction computing algorithm illustrated with an example to highlight the efficiency of the algorithm and then the top-N rewritings algorithm is provided.

Query Satisfaction Computing Algorithm

The contribution of the algorithm is its ability to estimate the degree of satisfaction of a conjunctive query rewriting of a fuzzy conjunctive query according to the fuzzy predicates occurring in the query. It provides an approximation of the average satisfaction degree of the fuzzy predicates occurring in the query by a tuple of the conjunctive rewriting result and this based on the constructed histograms. The first module of the algorithm estimates the result size of each conjunctive query rewriting; the second module constructs for each rewriting, the histograms approximating the distribution in the rewriting result of values of each attribute with which a fuzzy predicate is associated. The construction process is based on the histograms maintained on these attributes in the sources participating in the query. The inputs of the algorithm are the conjunctive fuzzy query Q, the membership functions defining the fuzzy conditions used in Q and finally the histograms describing the distributions of values in the attributes of interest in the query referred sources.

Let's, first, define the notion of a rewriting of a query using views and then useful present notations used in the algorithm. Given a query Q and a set of view definitions V_1, …, V_n, a rewriting of the query using views is a query Q' that refers only to the views V_1, …, V_n [2].

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The notations used in the different modules of the algorithm are: T(S) which represents the cardinality of Subgoal S, H[i] is The number of values in the interval I_i, attsum is the summarized attribute, occ(S, X) is the number of copies of one value subgoal S has in attribute X, Att("a") is the attribute containing the constant "a", V (S, X) is the number of distinct values (known or estimated) a Subgoal S has in attribute X, and μ (x) which is the membership degree of a value x to a fuzzy set.

The estimation of the rewriting result size is based on the histograms maintained for the join attributes in the relations participating to the query. The estimation process is described by the following module. See Figure 1.

The second module of the algorithm is devoted to construct the histograms, H_{res}, describing the distributions of values in the attributes to which fuzzy conditions are related. The construction process is described by the following part of the algorithm. See Figure 2.

The last module permits to compute the global satisfaction degree of a conjunctive query rewriting. Having available the histograms constructed by the module 2, the idea to compute the global degree is to determine the average satisfaction value of each fuzzy predicate in the query by a tuple returned by the rewriting, and then the weighted mean of these values in order to have a global vision of the rewriting. In this work, we propose to sum the average satisfaction degrees of the different fuzzy conditions and then to divide the sum by the number of fuzzy conditions or histograms. What we obtain, then, is an approximation of the average satisfaction degree of the conjunctive fuzzy query by a tuple returned by the rewriting. The computation of the global degree is given by the following module. See Figure 3.

Figure 1. Query result size estimation

```
1.  For each Subgoal S in Q' do
2.    T init (S) ⟵ T(S)                    //   T init (S) is the initial size of S before the equi-join
3.    occ (S, attsum) ⟵ T(S)/V(S, attsum)
4.    If Subgoal S is Negated Subgoal then
5.      For each parameter y do
6.        If y is a constant "a" then
7.          T(S) ⟵ T(S) / V(S, Att ("a")) × V(S, Att ("a")) – 1                    (2)
8.        End If
9.      End For
10.       Else
11.       If S is Arithmetic Comparison Subgoal then  // we consider T(≤) = T(<) & T(≥) = T(>)
12.         For each Subgoal S' in Q' do                    //   S' is not a Comparison Subgoal
13.           If belong(S', V) then     //   is a function that returns true if variable V belongs to the
                                          Subgoal S'
14.         If (S is Lower Comparison Subgoal)          //   S is a Comparison Predicate like <
              Or (S is Upper Comparison Subgoal) then   //   S is a Comparison Subgoal like >
15.           T(S') ⟵ ("a" – min(V) + 1) / (max(V) – min(V) + 1) x T(S)              (3)
16.       End If
17.         End If
18.       End For
19.       Else
20.         For each parameter y do
21.           If y is a constant "a" then
22.             T(S) ⟵ T(S) / V(S, Att ("a"))                                      (4)
23.           End If
24.         End For
25.       End If
26.     End For

27.     For each subgoal S in Q' do     //   S is not a Comparison Subgoal nor a Fuzzy Predicate
28.       If T init (S) ≠ T(S) then          //   T init (S) is S's size before selection operation
29.       For i from 1 to m do                //   m is the number of intervals
30.         If H[i] < occ (S, Attsum) then
31.           V (Ii, attsum) ⟵ H[i] / occ (S, attsum)
32.         Else
33.           H[i] ⟵ T(S) / V(S, attsum)
34.           V (Ii, attsum) ⟵ H[i]
35.         End If
36.       End For
37.     End If
38.   End For
39.   D ⟵ 1; N ⟵ 1
40.   Add (T, T (Subgoal 1)
41.   Size (T) ⟵ 1
42.   For i from 1 to nbSubgoal – 1 do
43.     For each variable y in Subgoal i do
44.       max ⟵ V(Subgoal i, y)
45.       join ⟵ False
46.       For j from i + 1 to nbSubgoal do
47.         If belong (Subgoal j, y)) then
48.           If Notbelong (T, Subgoal j)) then
49.             Add (T, T (Subgoal j))
50.             Size(T) ⟵ Size(T) +1
51.           End If
52.           max ⟵ max (max, V (Subgoal j, y))
53.           join ⟵ true
54.         End If
55.       End For
56.       If join then
57.         D ⟵ D × max
58.         If Notbelong (T, Subgoal i)) then
59.           Add (T, T (Subgoal i))
60.           Size (T) ⟵ Size (T) +1
61.         End If
62.       End If
63.     End For
64.   End For

65.   For i from 1 to Size (T) do
66.     N ⟵ N × T[i]
67.   End For
68.   T (res) ⟵ N / D
69.   Empty (T)
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Histogram construction

```
1. For each Fuzzy Predicate P in Q' do      //  P is a Fuzzy condition related to attribute attsum
2.    For each Subgoal S in Q' do
3.       If belong(S, attsum) then
4.            Add (T2, T(S))
5.       End If
6.    End For
7.    min ⟵ minimalSize (T2)
8.    If min = T (res) then
9.     hist_res (attsum) ⟵ hist_S (attsum)          //  S is the Subgoal with the minimal size
10.   Else
11.    For i from 1 to m do
10.      If H[i] < occ (S, Attsum) then
11.        V (Ii, attsum) ⟵H[i]
12.       Else
15.          V (Ii, attsum) ⟵ H[i] / occ (S, attsum)
16.          H_res[i] ⟵ T (res) × V (Ii, attsum) / V(S, attsum)
17.       End If
18.    End For
19.   End If
20. End For
```

Figure 3. Satisfaction degree computation

```
1. somme ⟵ 0
2. For each Histogram do
3.    μ ⟵ 0
4.    For k from 1 to m do
5.       μ_k ⟵ (μ_P(I_k.begin) + μ_P(I_k.end))/2
6.       μ ⟵ μ + μ_k × H[k]
7.    End For
8.    μ ⟵ μ / T(res)
9.    somme ⟵ somme + μ
10. End For
11. μ_moy (Q') ⟵ somme / Nbhist
12. Add (TAV, μ_moy (Q'))      //  TAV is a table containing for each rewriting
                               Q' the approximation of its average satisfaction
                               degree of the fuzzy predicates expressed in the
                               user query
```

Running Example

Let's consider a set of hospitals providing healthcare services where each hospital maintains a database accessible to the public through the Web. Each database lists all the physicians distributed among the different specialties in the corresponding hospital and provides information about their consultations with patients. Let's consider also a medical conjunctive fuzzy query submitted to a mediated system in-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tegrating the heterogeneous hospital databases. Each physician has a name and a fee which is the amount of a consultation. The consultation relation lists all the consultations being insured by physicians. An evaluation value is a score, belonging to the range [0, 20], associated to a consultation assured by a given physician in a given date. The hospital database schema is the following. See Figure 4.

The mediated schema exposed to the user is the hospital schema except that the relations Consultation and Physician have an additional attribute "hosp" identifying the hospital at which a consultation is being insured. The mediated schema is described below in Figure 5.

To illustrate the solution, suppose that we have the following two data sources. The first source s_1 lists all the physicians working in a hospital and their fees. This source is described in the mediator by the following view definition in Figure 6.

The second source s_2 lists among the physicians provided by source s_1 those assuring consultations with their evaluations; s_2 is described by the following view definition in Figure 7.

If a user needs to know the physicians having not expensive fees and having good evaluations, he submits to the data integration system the following query in Figure 8.

The data integration system is able to answer the previous query by joining the two sources s_1 and s_2. These sources are determined after reformulating the query, submitted over the mediated schema, into one that uses the crisp views v_1 and v_2 because v_1 (respectively v_2) mentions the relation Physician (respectively Consultation) which is revealed by the query and also selects the attribute name (respectively evaluation) which is also selected by the query. For more information about usability of a view to answer a query see (Halevy, 2001)). This query is then described in the following Figure 9.

Nonetheless, the system doesn't have any information about the quality of the result returned by the previous query. Consequently, the data integration system needs a more efficient way to answer a fuzzy query submitted in terms of its mediated schema; hence the idea of querying the integrated data sources through their summaries.

Let H1, respectively H2 be histograms describing the distribution of values in the attribute price in s_1 respectively in the attribute evaluation in s_2. The evaluation values described by a histogram are

Figure 4. Hospital schema

```
Patient (patinet_id, name, age)
Physician (physician_id, name, speciality, price)
Consultation (physician_id, patient_id, date, evaluation).
```

Figure 5. Mediated schema exposed to the users

```
Physician (physician_id, name, speciality, price, hosp)
Consultation (physician_id, patient_id, date, evaluation, hosp).
```

Figure 6. s_1 description using a view over the mediated schema

```
Create view v1 as
Select      physician_id, name, speciality, price
From        Physician.
```

Figure 7. s_2 description using a view over the mediated schema

```
Create view v2 as
Select      physician_id, hosp, date, evaluation
From        Consultation.
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Example of a medical conjunctive fuzzy query

```
Select    name, evaluation, hosp
From      Physician, Consultation
Where     Physician.physician_id =Physician.physician_id
          AND speciality = 'cardiology' AND price is
          NOT EXPENSIVE and evaluation is GOOD.
```

Figure 9. Example of a conjunctive fuzzy query rewriting

```
Select    name, evaluation
From      v1, v2
Where     v1.physician_id = v2.physician_id AND
          speciality ='cardiology' AND price is NOT
          EXPENSIVE and evaluation is GOOD.
```

grouped by physician_id and aggregated on average, so that each physician must have a single evaluation value. H1is made of the intervals I_1= [0, 100], I_2 = [101, 200], I_3 = [201, 300]. We have: H1 [1] = 10, H1 [2] = 25 and H1 [3] = 15. H2 is made of the intervals I1= [0, 10], I2 = [11, 20]. We have: H2 [1] = 15 and H2 [2] = 35.

We estimate, according to module 1, the size of the resulting relation based on the histograms maintained on s_1 and s_2. First, we estimate the size of the selection condition based on the histogram H maintained on the attribute Specialty as:

$$T(\sigma_{\text{Speciality="cardiology"}}) = H[I_k] / V(I_k, \text{Speciality}) \quad (5)$$

$$= 50 / 10 = 5$$

To estimate the final query result size, we need to estimate the size of the result of joining the selected tuples with those in the relation s_2. Here we assume that the same partition of the domain of physician_id is organized into fixed sized intervals $I_1, \ldots I_m$ in the two sources s_1 and s_2. The idea is to compute the result size of joining the values of each interval $I_{k\,(k=1 \ldots m)}$ of Hs_1 with those belonging to the interval I_k of Hs_2, and then the different sizes are summed in order to obtain the size of the query result.

$$T(Q) = \sum_{k=1}^{m}(Hs1 \,.\, \sum_{k=1}^{m}(Hs1\, H_{s1} \sum_{k=1}^{m}(Hs1\,.\,[] * T(s1) / T(\sigma_{\text{speciality="cardiology"}})) * H_{s2}[I_k] / \max(V(I_k, s1(\text{physician_id})), V(I_k, s2(\text{physician_id}))) \quad (6)$$

Let $H1_{res}$ and $H2_{res}$ be the histograms describing respectively the distribution of the fees and the evaluation values in the query result. Based on module 2, $H1_{res}$ and $H2_{res}$ are computed as follows.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$H1_{res}\left[1\right]=\frac{\left(V\left(I_1,\text{price}\right)*\text{query result size}\right)}{\sum_{k=1}^{3}V\left(I_k,\text{price}\right)}=\left(2*5\right)\Big/\left(2+3+5\right)=10/10=1 \tag{7}$$

$$H1_{res}\left[2\right]=\frac{\left(V\left(I_2,\text{price}\right)*\text{query result size}\right)}{\sum_{k=1}^{3}V\left(I_k,\text{price}\right)}=\left(3*5\right)\Big/\left(2+3+5\right)=15/10=1.5\approx 2 \tag{8}$$

$$H1_{res}\left[3\right]=\frac{\left(V\left(I_3,\text{price}\right)*\text{query result size}\right)}{\sum_{k=1}^{3}V\left(I_k,\text{price}\right)}=\left(5*5\right)\Big/\left(2+3+5\right)=25/10=2.5\approx 3 \tag{9}$$

$$H2_{res}\left[1\right]=\frac{\left(V\left(I_1,\text{evaluation}\right)*\text{query result size}\right)}{\sum_{k=1}^{2}V\left(I_k,\text{evaluation}\right)} \tag{10}$$

$$H2_{res}\left[2\right]=\frac{\left(V\left(I_2,\text{evaluation}\right)*\text{query result size}\right)}{\sum_{k=1}^{2}V\left(I_k,\text{evaluation}\right)} \tag{11}$$

Now having available the desired histograms for the different query rewritings enables to determine how closely each rewriting matches the fuzzy conditions specified in the user query. Let us consider the two fuzzy sets "NOT EXPENSIVE" and "GOOD" defined by the following fuzzy membership functions in Figures 10 and 11.

The module 3 of the algorithm provides the average satisfaction degrees respectively of the two fuzzy predicates "NOT EXPENSIVE" and "GOOD"**:**

((1 * 1) + (0.495 * 2) + (0 * 3))/6 = 1.99/6 ≈ 0.33

((0.35 * 2) + (0.89 * 3))/5 = 3.37/5 ≈ 0.67.

The actual average satisfaction degrees respectively of the fuzzy predicates "NOT EXPENSIVE" and "GOOD" equals 1.5/5 = 0.3 and 4.14/5 ≈ 0.82. In this particular case, the degrees computed (0.33) and (0.67) are relatively good approximations.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Membership function defining the fuzzy set price

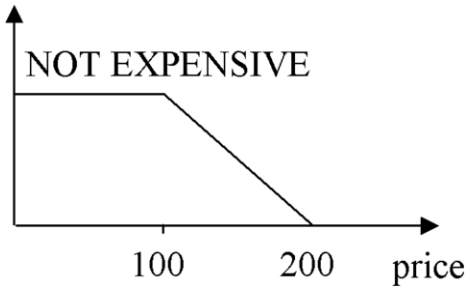
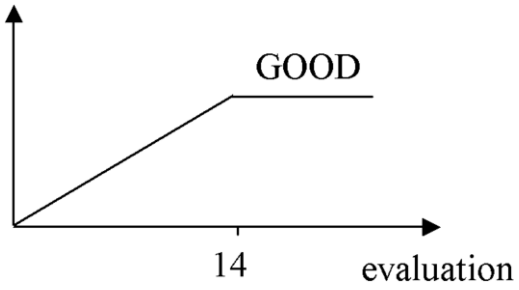


Figure 11. Membership function defining the fuzzy set evaluation



Top-N Rewritings Solution

Different techniques have been proposed in the literature to determine the best query rewritings that are likely to return the best answers in the context of answering fuzzy query using views. In (Jaudoin & Colomb & Pivert, 2009), authors deal with the problem of answering query through probable answers instead of certain answers. The authors propose an algorithm for computing the top-K rewritings where a score between 0 and 1 is attached to each rewriting Q' of a query Q, which corresponds to the probability for a tuple returned by Q' to be an answer to Q. The authors of (HadjAli & Pivert, 2008) deal with fuzzy query answering using views where the view descriptions may include fuzzy predicates. The proposed approach associates a satisfaction threshold to each query and the best views, those which provide answers whose satisfaction degree is over the threshold specified by the user, are determined through a matching of the fuzzy query against the fuzzy views. In the same context where the views are fuzzy, the paper (Jaudoin & Pivert, 2011) constitutes a contribution to the definition of flexible Local-As-View mediation systems. It proposes an approach to rewriting fuzzy queries using imprecise views where the query and the views may include fuzzy value constraints. These constraints describe attribute values as a set of elements attached with a degree in [0, 1] that expresses the plausibility attached to a given element while in the queries, they denote preferences.

The essence of Top-N rewritings algorithm is to generate the N best rewritings for a fuzzy query Q without computing its possible rewritings. It tries to achieve significant reduction in query execution time as it considers minimal number of rewritings that are likely necessary to return the Top-K tuples for a submitted medical fuzzy query. It generates the best rewriting fitting user preference first, then the next best rewriting second until having the N required rewritings. We mention that the numbers (N) of the rewritings to find and (K) of the desired tuples are different and N is an internal variable fixed by the middleware depending on the K value. Our algorithm is the first study to generate the Top-N ordered rewritings without computing them or searching all possible ones for a medical fuzzy query. Its originality is for searching the N best rewritings, it considers only N rewritings and it is certain that are the best ones. Top-N rewritings algorithm is inspired from the MiniCon (Pottinger & Halevy, 2000) one generating the maximally contained rewriting of a conjunctive query Q. MiniCon provides a maximally contained rewriting for Q and not an equivalent one because in virtual information integration systems, it is assumed that the sources are incomplete including only a subset of answers and the views defined over the global schema may not refer to all databases stored information (Open World Assumption). MiniCon can be used as part in our algorithm after modifying it. The first modification concerns the input and the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

output of the algorithm. The input of Top-K rewritings algorithm is a fuzzy conjunctive query, a set of crisp views ordered by their satisfaction degrees for a fuzzy query and the number N, and the output is the N best rewritings according to the fuzzy medical submitted query. The second modification is in the MCD formation step. In fact, this step is not done for all views at the algorithm beginning.

Nevertheless, it is performed for the most important view in the views' list then for the second most important view, etc. The higher view satisfaction degree is, the most important the view is for a rewriting. The second step of the MiniCon algorithm consisting in combining the MCDs is also modified. In fact, Top-N rewritings algorithm searches the best rewritings using the MCDs of the most important view, then the second most important rewritings by combining the MCDs of the most important view and the MCDs of the second most important view, etc. A rewriting importance is evaluated by a monotonic satisfaction function aggregating the satisfaction of the different views composing the rewriting.

Our algorithm stops once it reaches the N needed rewritings.

MiniCon Algorithm

Preleminaries

- **Queries Containment**: A query Q_1 is contained in a query Q_2 if for each database instance D, the result of Q_1 applied to D is included in the result of Q_2 applied to the same instance D (Halevy, 2001).
  - Formally: $Q1 \subset Q2$ *if* $Q1\ (D) \subset Q2\ (D)$.
- **Mapping**: A mapping between two queries Q_1 and Q_2 is a mapping between the variables of Q_1 and the variables of Q_2. To form a mapping between two queries, the following necessary conditions must be satisfied:
  - A variable must not be mapped at the same time to different variables or constants.
  - A constant must not be mapped to a variable or a constant other than itself.
- **Containment Mapping**: Containment mapping is a necessary and sufficient condition to check the containment between queries (Halevy, 2001). This means that to conclude the inclusion between a query Q_2 and a query Q_1, it is sufficient to prove that there is a containment mapping between them without executing them. A mapping (t) is a containment mapping if it maps every sub-goal of Q_2 to a sub-goal of Q_1 and the head of Q_2 to the head of Q_1.
- **Queries Equivalence**: A query Q_1 is equivalent to a query Q_2 if Q_2 is included in Q_1 and Q_1 is included in Q_2.
  - Formally: $Q_1 \equiv Q_2$ *if* $Q_1\ (D) \subseteq Q_2\ (D)$ *and* $Q_2\ (D) \subseteq Q_1\ (D)$.
- **Equivalent Rewritings**: Let A be a schema, $V = \{V_1, V_2, \ldots, V_n\}$ a set of views over A and Q a query over A.
  - Q' is an equivalent conjunctive rewriting (Halevy, 2001) for a conjunctive query
  - Q if:
  - Q' refers only to the views of V.
  - Q' is equivalent to Q.
- **Maximally Contained Rewriting**: A maximally contained rewriting aims at finding as many as possible answers from the provided views. Q' is a maximally contained rewriting for a conjunctive query Q if:
  - Q' refers only to the views of V.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Q' is included in Q.
- There is no other rewriting Q_1 such as $Q' \subset Q1 \subset Q$ and $Q' \neq Q_1$.

MiniCon Algorithm

MiniCon algorithm (Pottinger & Halevy, 2001) is proposed in the Tukwila project. It is used to search a maximally contained rewriting for a classic conjunctive query Q using a set of conjunctive views V. It cares about the interaction between the query variables of joined predicates. In fact, if a view contains a target sub-goal participating in a join in the query, the join variables should appear in the considered view head in order to join it with other views. Minicon (see Figure 12) begins by considering each query sub-goal in isolation. Once it finds a partial mapping between a view sub-goal g and a query sub-goal g', it looks at the query variables of the join predicates to find the minimal additional set of the query subgoals that must be mapped to the view sub-goals in order V be usable. This set of query sub-goals that have to be mapped to sub-goals of a view and the related mapping information is called MiniCon Descriptions (MCDs). It says how the query sub-goals can be covered by a view. It is considered as a fragment of the containment mapping from the query to its rewritings.

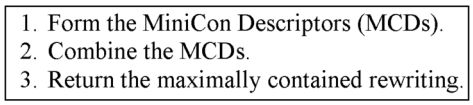
1. Form the MiniCon Descriptors (MCDs).
2. Combine the MCDs.
3. Return the maximally contained rewriting.

The following conditions have to be respected when forming the MCDs:

- **Condition 1**: Every query distinguished variable or query constant must be mapped to a view distinguished variable.
- **Condition 2**: If a variable X is a part of a join predicate which is not included in a view V then it must appear as a head variable in V to join other views in a query rewriting.

After determining the MCDs of the different views, they are combined to form query rewritings while respecting the next safe condition: rewritings must contain disjoint MCDs covering all the query sub-goals. There is no need to check the containment test as all the formed rewritings are valid ones thanks to the MCDs construction conditions. The maximally contained rewriting is the union of the different valid rewritings from which views' redundancies are removed.

Figure 12. MiniCon algorithm steps



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Top-N Rewritings Algorithm

The input of the Top-N rewritings algorithm (see Figure 13) is a sorted list of crips views, the number of rewritings N and the fuzzy conjunctive query Q. The sorted crisp views (V_i,…, V_n) are examined according to their orders in the list (see the following running example): the best crisp view is used first then the second until obtaining the N needed rewritings. When a crisp view V_i is chosen from the ordered list of crisp views, the algorithm looks for the MCDs that can be formed with to cover the query sub-goals. Next, the algorithm searches the rewritings that can be formed using only the MCDs of Vi, then the rewritings that can be formed using the MCDs of V_i and the MCDs of the previous examined crisp views from V_1 to V_{i-1} (if these MCDs exist). The rewriting research is facilitated by the monotonic propriety of the satisfaction function. In fact, the research begins with the rewritings formed by one MCD then by two MCDs then by three MCDs, etc. The limit number of MCDs in a rewriting depends on the number of the previous observed views. Our algorithm abandons the researches ones it reaches the number N of rewritings. The found rewritings are the best that can be searched. See Figure 13.

Figure 13. Top-N rewritings algorithm

```
Input
 Q: a conjunctive query
 N: the number of rewritings
 V: a list of sorted views of cardinality n
Output
 Result: Top-N rewritings
Begin
1. Result ◄— Ø
2. NbRewriting ◄— 0
3. NbViews ◄— 1
4. While(NbRewriting < N)and(NbViews < = n)
5. Begin
6.    MCD [NbViews] ◄— createMCD(V[NbViews], Q)
7.    NbVR ◄— 1
8.    While(NbRewriting N)and(NbVR<=nbsubgoal) and(NbVR <=NbViews)
9.    Begin
10.    rewritings ◄— findRewritings(MCD[NbViews], NbVR, Q)
11.    If (card (rewritings) ≠ 0) Then
12.       x ◄— 1
13.       While (x<=card(rewritings))and
14.       (NbRewriting <N)
15.       Begin
16.        Result ◄— add (rewritings[x], Result)
17.        x ◄— x+1
18.        NbRewriting ◄— NbRewriting+1
19.       End
20.    End if
21.    NbVR ◄— NbVR+1
22.  End
23.  NbViews ◄— NbViews+1
24.End
25.Display (Result)
26.End
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Running Example

Assume we have four crisp views (from V_1 to V_4) to rewrite a medical fuzzy query Q. Each view has a satisfaction degree (F (V_i)) such that F (V_i) is higher than F (V_{i+1}). Assuming also that each combination of crisp views is a rewriting for Q. Using Top-N rewritings algorithm (Table 1), the rewritings are generated from the most to the least important (from R_1 to R_{14}): R_1 is the best rewriting as it is formed by the view V_1 having the higher view satisfaction degree for Q. Then R_2 is generated having the degree of V_2 then R_3, etc. The generation of R_2 is done before R_3 because a rewriting having the least number of views is the most important in the case of equal satisfaction degrees. Top-N rewritings algorithm stops once it reaches the number N of rewritings. If N is 1, it stops once R_1 is found without exploring all the research space as R_1 is the best rewritings' research. When it finds R_2, it returns R_1 and R_2, etc.

Top-k Tuples Generation

In recent years, many research works have acknowledged the need for flexible ways to access database information. A typical example is the top-K query where the system aims at providing the user with the best K answers. The most relevant research works studied the problem of efficient answering of top-K queries using views (Das & Gunopulos & Koudas & Tsirogiannis, 2006, Xie & Lakshmanan & Wood, 2013).

In (Das & Gunopulos & Koudas, 2006), authors deal with the problem of answering top-K queries using data for example in conjunction from a collection of views where a view in this context is defined as a materialized form of a top-*K* query asked previously according to some monotone combining function defined on a subset of the attributes of a relation. The main contribution is to propose an efficient algorithm called LPTA to identify, among a collection of views, the most promising views to use for query answering.

However, the performance of LPTA suffers because of iterative calls to a linear programming subprocedure, which can be especially problematic when the number of views is large or if the dimensionality of the dataset is high. To improve the efficiency of the LPTA algorithm, authors in (Xie & Lakshmanan

Table 1. Ranked rewritings generation

Query Rewritings	View Satisfaction Degree
$\mathbf{R_1}$:-V_1	F(V_1) >
$\mathbf{R_3}$:-V_1, V_2 $\mathbf{R_2}$:-V_2	F(V_2) >
$\mathbf{R_7}$:-V_1,V_2, V_3 $\mathbf{R_6}$:-V_1, V_3 $\mathbf{R_5}$:-V_2, V_3 $\mathbf{R_4}$:-V_3	F(V_3) >
$\mathbf{R_{15}}$:-V_1,V_2, V_3, V_4 $\mathbf{R_{14}}$:-V_1, V_3, V_4 $\mathbf{R_{13}}$:-V_2, V_3, V_4 $\mathbf{R_{12}}$:-V_1, V_2, V_4 $\mathbf{R_{11}}$:-V_3, V_4 $\mathbf{R_{10}}$:-V_1, V_4 $\mathbf{R_9}$:-V_2, V_4 $\mathbf{R_8}$:-V_4	F(V_4)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

& Wood, 2013) propose to take advantages of the cached results of previous queries. They propose the LPTA$^+$ algorithm to efficiently answering top-k queries using cached views. The experiment results, studying the performance of the LPTA$^+$ algorithm, show the interesting reduction of waiting time to provide the user with the top K answers.

In this section, we consider the problem of how to efficiently exploit the output of the Top-N rewritings algorithm to retrieve the top-K tuples. Each rewriting of the Top-N ones satisfies in general the fuzzy query but it doesn't affirm that all its tuples are better than those of the succeeding rewriting. Hence, two possible solutions can be envisaged to find Top-K tuples of Q using its top-N rewritings which are:

- Running the Top-N rewritings, sorting the resulting tuples and giving users only the K best ones. If a user wants the second best K tuples, he can require them. This solution gives the best tuples for user query but the disadvantage consists in the running time of the all Top-N rewritings.
- Running the first rewriting from the Top-N ones. If it is resulting tuples cardinality is equal or over than K, the tuples are sorted and only the first K of them are returned for users. Else, the succeeding rewritings are computed one by one until having K tuples to return. This solution helps users quickly obtaining globally satisfying answers for their fuzzy queries. But, they aren't sure that the received answers are the best ones.

Query Cache

To enhance the fuzzy query processing, we suggest a cache memory in the middleware storing the last implicit queries' results triggered by the physicians' PDAs. Same predefined implicit queries can be submitted by different PDAs located in a same region. Consequently, there is no need to re-process them as their Top-K results are readily available in the cache. Cache verification is performed for each submission of an implicit query. If it doesn't exist in the cache, the entire query answering process described in the previous sections will be executed and its result will be stored in the cache else, the answer will be retuned instantly. The cache is initialized for each period of time to have actualized information.

FEASABILITY OF THE STUDY

The system architecture is simple with no complex communication architecture. On one hand, users, that can be either patients or physicians or other healthcare workers, are familiar with the use of mobile devices such as mobile phones. On the other hand, the setup of the bio-sensors is somewhat complex, but the advantages taken from the use of these bio-sensors in the context of improving healthcare services have much more importance comparing to the setup complexity of these bio-sensors. The implementation of the system is not expensive. The bio-sensors are generally low cost and the communication is also cheap due to the use of low cost wireless adapters and low cost mobile devices. The security is a challenging issue which concerns the safety of the communicated information transmitted between the different components. This information should be protected from malicious attacks and security vulnerability. In their turn, the communication protocols are flexible and can be adjusted to the communicated components which can be either personal computers or Cell phones or PDAs, etc.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

The acknowledged integration of fuzzy logic, in different domains, arises the difficulty of estimating the pertinence of information sources for conjunctive fuzzy queries which is a very challenging problem. However, the introduction of fuzzy logic in the medical domain is very promising in the sense that users can use natural language and vague terms when querying medical applications and efficiently obtain best answers in the shortest waiting time.

In the literature, there are no algorithms proposed in middleware environments to determine the pertinent query rewritings likely to return the best answers for a conjunctive fuzzy query. Therefore, implementing such an algorithm will be very helpful for medical applications users, as we illustrated with a simple case study of a healthcare application.

In this paper, we presented a two steps method destined to find as quickly as possible and efficiently the best satisfying answers for a medical fuzzy query. In the first step, we presented the Top-N rewritings algorithm, a novel algorithm for generating ordered top-N rewritings for a fuzzy medical query. At the best of our knowledge, this algorithm is the first one to generate the N best rewritings without computing or searching all possible ones. N is a variable depending on the K desired tuples considered as best answers. In the second step, we proposed two different solutions on how exploiting the output of the Top-N rewritings algorithm in order to recover the Top-K tuples.

Finally, we briefly discussed the use of a query memory cache in order to avoid useless query processing when it is already available in the cache.

The proposed Top-N rewritings algorithm is based on the hypothesis of the uniform distribution to estimate the pertinence of information sources. As a future work, we plan to improve techniques of estimating the pertinence of information sources in order to obtain reliable results based on more accurate estimates.

REFERENCES

Ahamed, S. I., Haque, M. M., Stamm, K., & Khan, J. (2007). Wellness assistant: A virtual wellness assistant using pervasive computing. *Proceedings of the ACM Symposium on Applied Computing (SAC)* (pp. 782-787). 10.1145/1244002.1244177

Bardram, J. E. (2004). Applications of context-aware computing in hospital work: Examples and design principles. *Proc. of the ACM Symposium on Applied Computing* (pp. 1574–1579).

Bardram, J. E., Christensen, H. B., & Olsen, A. K. (2004). *Hypernavigation in the physical space: Adapting presentations to the user and to the situational context (Tech. Rep.)*. Centre for Pervasive Computing.

Bardram, J. E., Hansen, T. R., Mogensen, M., & Sogaard, M. (2006). Experiences from real-world deployment of context-aware technologies in a hospital environment. *Ubicomp*, *6*, 369–386.

Bartolomeu, P., Fonseca, J., & Vasques, F. (2008). Challenges in health smart homes. *ATDMCP – Workshop on Ambient Technologies for Diagnosis and Monitoring Chronic Patients, part of the 2nd International Conference on Pervasive Computing Technologies for Healthcare.*

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Black, J. P., Segmuller, W., & Cohen, N. (1995). Pervasive computing in health care: Smart spaces and enterprise information systems. *Proceedings of the IEEE International Conference on Communications* (pp. 3–8).

Boers, N. M., Chodos, D., Huang, J., Gburzy'nski, P., Nikolaidis, I., & Stroulia, E. (2009). The Smart Condo: visualizing independent living environments in a virtual world. *Pervasive Health. Proceedings of the 3rd International Conference on Pervasive Computing Technologies for Healthcare*. IEEE. 10.4108/ICST.PERVASIVEHEALTH2009.6020

Bosc, P., Hadjali, A., Jaudoin, H., & Pivert, O. (2007). Flexible querying of multiple data sources through fuzzy summaries. *Proc. of the FlexDBIST Conference*, Regensburg, Germany.

Christensen, H. B., & Bardram, J. (2002). Supporting human activities – exploring activity-centered computing. *Proc. 4th of the International Conference on Ubiquitous Computing* (pp. 107-116). Springer-Verlag.

Das, G., Gunopulos, D., Koudas, N., & Tsirogiannis, D. (2006). Answering top-k queries using views. Proceedings of VLDB, Seoul, Korea (pp. 451-462).

Dushka, O., Genesereth, M., & Levy, A. (2000). Recursive query plans for data integration. *Journal of Logic Programming*. *Special Issue on Logic Based Heterogeneous Information System*, *43*(1), 49–73.

Favela, J., Rodriguez, M., Preciado, A., & Gonzalez, V. M. (2004). Integrating context-aware public displays into a mobile hospital information system. *IEEE Transactions on Information Technology in Biomedicine*, *8*(3), 279–286. doi:10.1109/TITB.2004.834391 PMID:15484433

Gaglio, S., Re, G. L., Martorella, G., Peri, D., & Vassallo, S. D. (2014). Development of an IoT environmental monitoring application with a novel middleware for resource constrained devices. *Proceedings of the 2nd Conference on Mobile and Information Technologies in Medicine (MobileMed'14)*.

Hadjali. A., & Pivert, O. (2008). Towards fuzzy query answering using fuzzy views - A graded-subsumption-based approach. Proceedings of ISMIS, Toronto, Canada (pp. 268-277).

Halevy, A. Y. (2001). Answering queries using views: A survey. *Very Large Data Bases Journal*, *10*(4), 270–294. doi:10.1007007780100054

Halteren, A. V., Bults, R., Wac, K., Dokovsky, N., Koprinkov, G., Widya, I., ... Herzog, R. (2004). Wireless body area networks for healthcare: The mobihealth project. *Studies in Health Technology and Informatics*, *108*, 181–193. PMID:15718645

Jaudoin, H., Colomb, P., & Pivert, O. (2009). Ranking approximate query rewritings based on views. Proceedings of FQAS, Roskilde, Denmark (pp. 13-24).

Jaudoin, H., & Pivert, O. (2011). Rewriting fuzzy queries using imprecise views. Proceedings of ADBIS, Vienna, Austria (pp. 257-270).

Labbadi, W., & Akaichi, J. (2010). Estimation of information sources pertinence in mediated data integration systems. *Proc. of the 3rd. International Conference on Information Systems and Economic Intelligence,* Sousse, Tunisia (pp. 157-288).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lenzerini, M. (2002). Data integration: A theoretical perspective. *Proc. of the ACM Symposium on Database Systems* (pp. 233–246).

Levy, A. Y., Rajaraman, A., & Ordille, J. J. (1996). Querying heterogeneous information sources using source description. *Proc. of the 22nd International Conference on Very Large Data Bases*. Mumbai (Bombay).

Luigi, A., Antonio, I., & Giacomo, M. (2010). The Internet of Things: A survey. *Computer Networks*, *54*(15), 2787–2805. doi:10.1016/j.comnet.2010.05.010

Nitin, R. K., & Sujata, T., Dr. (2015). Pervasive monitoring of m-health care using android. *International Journal of Science and Research*, *4*(4).

Poosala, V., Ioannidis, Y., Haas, P., & Shekita, E. (1996). Improved histograms for selectivity estimation of range predicates. Proc. of ACM SIGMOD (pp. 294–305). doi:10.1145/233269.233342

Pottinger, R., & Halevy, A. Y. (2001). MiniCon: A scalable algorithm for answering queries using views. *Very Large Data Bases Journal*, *10*, 182–198.

Shahriyar, R., Bari, M. F., Kundu, G., Ahamed, S. I., & Akbar, M. M. (2000). Intelligent mobile health monitoring system. *International Journal of Control and Automation*, *2*(3), 13–28.

Shoaib, M., Elbrandt, T., Dragon, R., & Ostermann, J. (2010). Altcare: safe living for elderly people. *Proceedings of the 4th International ICST Conference on Pervasive Computing Technologies for Healthcare*. 10.4108/ICST.PERVASIVEHEALTH2010.8878

Ullman, J. D. (2000). Information integration using logical views. *Theoretical Computer Science*, *239*(2), 189–210. doi:10.1016/S0304-3975(99)00219-4

Varshney, U. (2003). Pervasive healthcare. *IEEE Computer*, *36*(12), 138–140. doi:10.1109/MC.2003.1250897

Varshney, U. (2006). Patient monitoring using infrastructure-oriented wireless LANs. *International Journal of Electronic Healthcare*, *2*(2), 149. doi:10.1504/IJEH.2006.008829 PMID:18048241

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. [Special Issue on Pervasive Healthcare]. *Journal on Mobile Networks and Applications*, *12*(2-3), 111–228. doi:10.100711036-007-0018-0

Varshney, U. (2006). Using wireless technologies in healthcare. *International Journal on Mobile Communications*, *4*(3).

Xie, M., Lakshmanan, L., & Wood, P. (2013). Efficient top-k query answering using cached views. Proceedings of EDBT/ICDT, Genoa, Italy (pp. 489-500).

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 12(2); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 46-64; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 15
Development of Portable Medical Electronic Device for Infant Cry Recognition: A Primitive Experimental Study

Natarajan Sriraam
M. S. Ramaiah Institute of Technology, India

S. Tejaswini
M. S. Ramaiah Institute of Technology, India

Ankita Arun Chavan
M. S. Ramaiah Institute of Technology, India

ABSTRACT

Infant cry is considered as the first biological signal communicated by the newborns to the mother and the outside world. In NICU, there is a huge need to monitor the physiological conditions of the premature and full term infants. In such environment, the cry signal conveys significant information in terms of the distress of the infant and thereby needs of the infants can be well attended. Establishment of the portable device is thus essential for such scenario. This work enumerates a primitive experimental study on developing portable electronic medical device for infant cry recognition. The setup comprises of a voice recorder, NI My RIO processor with a Lab view followed by the LCD Tablet for display. The cry signals were recorded by placing the voice recorder in the incubator and were qualitatively assessed by the clinician and confirm its suitability for the experimental study. A threshold based classification was employed which sends the voice based alarm to the tablet. The continuous monitoring of such facility in NICU provides the behavioral status of the infant

DOI: 10.4018/978-1-7998-1204-3.ch015

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Infant cry, a biological signal that connects the mother and newborn is considered as valuable information for clinical diagnosis by the neonatologists (Kusaka et al, 2008; Bard et al., 2008; Chang et al.2016). The infant cry helps in recognizing the physical state of the baby such as pain, hunger, discomfort, wet diaper, fear, illness etc. In a typical neonatal intensive care unit (NICU), where pre-mature newborns & full term babies are undergoing treatment, their physical behavioral status needs to be known for any clinical decision. The cry signal reflecting the physical and psychological state of the newborns provides the opportunity for the clinical as well as mother to understand their needs. Though NICU'S combines advanced technology and trained healthcare professionals to provide specialized care, continuous monitoring of babies to cater the needs is a challenging task. It has been reported that the crying behavioral pattern of infants were closely associated with the maternal emotional distress during the post pattern period (Kusaka et al, 2008).

The analysis of the biological infant cry reveals the physical state of the babies including hunger, pain, fear, wet diaper etc. It is the audio frequency (250Hz to 600Hz) that distinguishes the normal cry from pathological cries. In the last two decades, attempts have been made to analyze the cry signals but no technology has been developed to monitor the pre-mature/full term babies in a NICU environment. Studies revealed that due to lack of facilities and continuous monitoring of the babies in the incubator and warmer, the mortality rate is exponentially growing high.

Further current clinical practice mainly relies on the subjective behavioral observation of infants in NICU. There is still lack of sophisticated monitoring system to screen and detect the patho-physiological changes of the infants.

This research work highlights a primitive experimental study on developing portable medical electronics device for infant cry recognition. The portable device comprises of three units, namely: voice recorder, real-time NI myRio processor and LCD tablets. The study simulates the typical NICU unit environment with recording of premature infants & full term babies in an incubator and warmer.

Literature on Infant Cries Detection

Literature works have been emphasized that analysis of infant cry an important field. This part briefly describes the recent methods for analysis of infant cry and pain assessment.

Dhanashri U.S. Talauliker and Nayana Shenvi (2015) have studied the characteristic of cry for preterm and full term infants. The cry characteristic reflects the development and possibly the integrity of central nervous system (CNS). They noticed that there is a difference between the cry patterns of preterm and infants with neuro-physiological conditions with fundamental frequency. Their aim was to preprocessing for eliminating silenced region of cry signal and estimates the fundamental frequency (pitch) using time & frequency domain analysis. The study was of main importance to exploring brain function at early stages of child development for timely diagnosis of neonatal disease malfunction.

D. Bard, P. Runefors, E. Arnbjornsson et al. (2008), discussed about the hypothesis that cry can be used as a tool to detect signs of nocioceptive pain. The pitch frequency is extracted from cry with time domain methods. The fluctuations of these were analyzed in filter. The comparison was done between the original and new results.

Sandhya and Premanand (2016) studied the mechanism of different types of cry. The dataset for the work consisted of 27 babies which represents of 7 hunger cries, 4 sleepy, 10 pain and 6 uncomfortable

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

cries. The studied also considered the Dunstan Baby Language (DBL) for the manual classification. The LFCC features were extracted and VQ codebooks was used to compress the feature vector. High quality results were obtained with LFCC compared to MFCC features.

Warlaumont et al.; (2010) have suggested features with self-organizing map and single-layer perception neural network as pattern classifier. Cries of six infants between 3 and 2 month were considered for the study. The data driver approach classifies the utterance into paralinguistic phonatory conditions.

Rautava et al.; (2006) studied the acoustic quality of cry in very low birth weight infants. The recordings were done after (MMR vaccination) at pediatric clinic and first cry utterance was studied using Praat software. Eleven dichotomist variables were evaluated from the cry pattern of each infant. The differences in the acoustic cry characteristics between the groups were also studied.

An automated infant cry recognition system was proposed to classify normal, deaf and asphyxiating infants (Galaviz & Reyers- Garcia.; 2004). Two features mel frequency cepstral coefficients and linear prediction co-efficient based acoustic features extracted were extracted and input delay neural network with gradient descent back propagation training algorithm was applied. A classification accuracy of 81% was shown.

Chang et al.: (2016) have proposed a sequential forward floating procedure to recognize the best features among time & frequency domain based features proposed for the study. A directed acyclic graph support vector machine was employed for the classification. Three variants hunger, pain, feeling sleepy were considered and an overall classification accuracy of 93.75% was achieved.

N. Sriraam & Tejaswini (2014) have done a pilot study on infant cry detection and pain scale assessment. The study highlights the critical need for the portable system for cries detection. A preliminary that makes use of mel frequency pattern with recurrent neural network classification was applied and the ROC analysis confirmed the efficiency of the classifier.

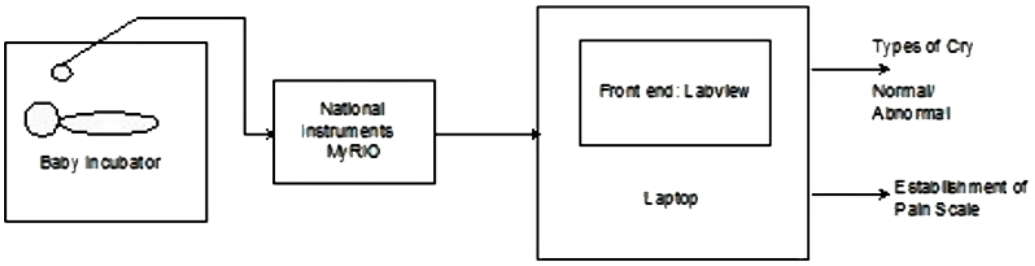
Varsharani & Sardar (2015) have applied linear frequency cepstrum coefficients and vector quantization code book for infants cry detection samples were collected from babies of less than six months old. A recognition rate of 91.58% was attained.

Conceptual Model

Figures 1 and 2 show the conceptual model for the proposed medical device for infant cry recognition considering the NICU environment. The model can be visualized a single baby incubator or multi incubator.

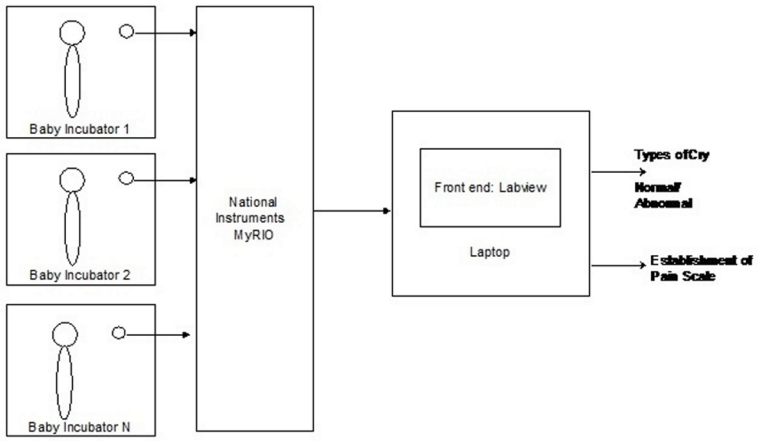
The voice recorder placed inside the incubator can be virtually connected to the real time processor, national Instrument NI My RIO processor. It has the figure potential to connect single and multi-voice

Figure 1. Proposed model for a signal baby incubator in NICU



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Proposed module for a multi incubator in NICU



recorders through its port facility. For initial study, wired mode is used. The my RIO processor is configured in such a manner that with a help of lab view software the raw recordings are converted into digitized data. The typical block diagram of infant cry system comprises of mainly data acquisition, preprocessing, feature extraction, feature reduction, classification and display.

Data Acquisition

The cry samples that have to be used for this study are collected from few healthy babies aged less than 6 months, few ill, premature infants. The gender is also being considered when recording the cry signals as the fundamental frequency of male infant is greater than female infant. The cry is recorded by using a digital voice recorder in context with babies cry during pain, hunger and fear. The voice recorder with sampling frequency 44.1 KHz with sampling resolution of 16 bits is used.

Preprocessing

The raw cry signal is preprocessed in order to eliminate noise in the recorded cry signal. Normalization, cleaning and filtering is done in order to eliminate noise in recorded signal and it should be suitable to extract the required parameters such as MFCC's, pitch and LPC etc.

Feature Extraction

Mel frequency cepstral coefficients are one of the common and well known acoustic features. MFCC's are used to encode the speech signals that contain irrelevant information with large amounts of storage space into compressed and useful features (A. Zabidi, et al., 2010). The characteristics of the cry can be attributed to the characteristics of vocal tract system and excitation source and supra segmental characteristics. The spectral peaks, bandwidths and slopes are unique for each condition of cry. Hence if we extract this information one can discriminate the infant cry condition (Ali Messaoud, ChakibTadj, 2011). Pitch is one of the important and distinguished acoustic features to analyze the infant cry. Linear

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

prediction analysis is a time domain analysis which has been used frequently in infant cry analysis. Linear prediction analysis attempts to predict a speech sample through linear combination of several previous samples.

Feature Reduction

The extracted features are further processed for acquiring the required signal cry signal. It is possible to apply filter banks of the Mel or Bark scale in order to perform pitch wrapping or subjective loudness warping to the cry segments prior to cry analysis. This in turn selects the extracted feature for classification

Classification

The infant cry signals have distinct patterns which can be recognized with pattern classifiers such as artificial neural network and support vector machine. For example, time delay neural network classifies cries between normal, deaf and asphyxiated infants. Feed forward neural network, recurrent neural network to differentiate between cries from three different stimulus situations such as pain, fear and hunger. Support vector machine using kernels like linear, quadratic, polynomial and radial basis function kernel will be applied to differentiate different types of cry.

Display

The display is the important part of the infant cry system. The processed, classified cry signal is distinguished and named accordingly is displayed on the laptop LCD screen. The display gives the information about the cry which is normal or abnormal. The condition of both the cry is displayed.

The cry of an infant is recorded using speaker which is interfaced to the National Instrument Data Acquisition System. The audio signal is amplified and processed using Labview software for detecting the type of infant cries. Once the optimal solution is arrived, then the proposed scheme will be extended for simultaneous recordings of cries from various babies. The proposed system thus will help to monitor the well-being of the infants in a typical neonatal ward.

Figure 3. Components of the proposed study

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Experimental Study

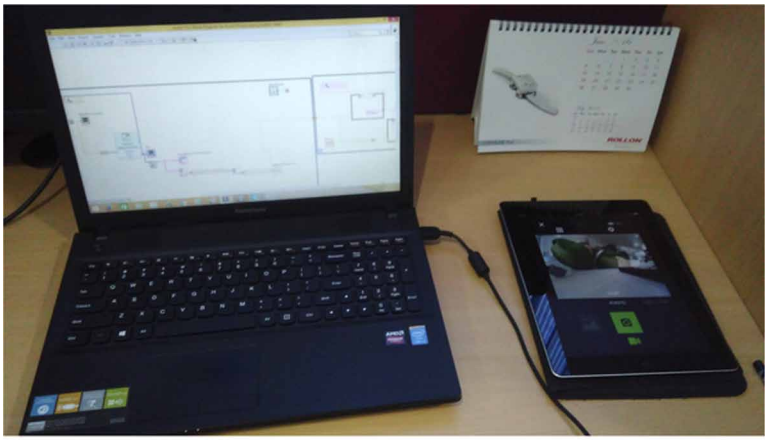
After due ethical clearance, the cry of the infant was recorded at M.S. Ramaiah Medical Hospital, Bangalore, India. 62 infants, 38 full-term and 24 pre- term babies were involved and the aim was to discriminate the cry and wet diaper. The proposed study aims to record, process and convey the status of the infant to the staff nurse, neonatologist in NICU. To execute this operation, Lab based experimental setup was arranged at the Centre for medical electronic and computing (CMEC):

The prerecorded voice recorder was placed in the incubator. Then the recorder was connected to myRio processor.

The final outcome is shown in the LCD tablet where the alarm will be heard if any pathology is detected. Further voice based signal will be played to indicated the nature of the infant cry.

Figure 4. Experimental setup arrangement

Figure 5. Workstation arrangement kept in different room



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

METHODOLOGY

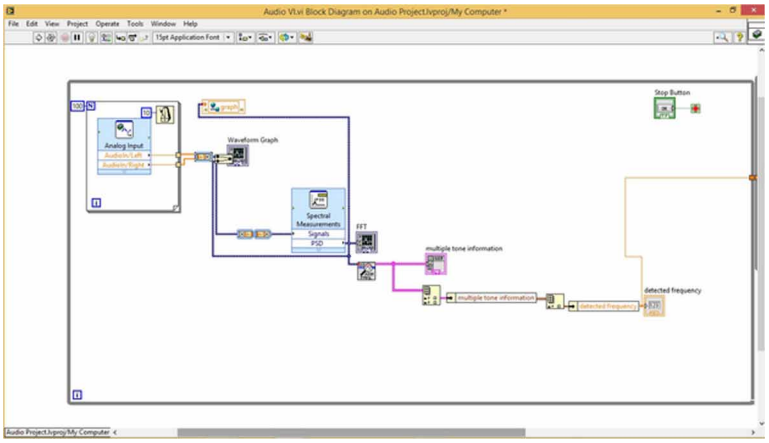
The focus of the primitive study is to understand the various aspects of infant cry analysis and classification. The work proposed implements automatic infant cry analysis which records infant cry which is a vocal signal and the signal is processed and analyzed to differentiate the healthy cries. The differentiated cries will further distinguish the physical state of a healthy infant such as hunger, fear, pain and wet diaper.

The infant cry detection deals mainly with the pitch which is the fundamental frequency from the analysis of infant cry. Infant cry analysis is a pattern recognition problem. It is similar to speech recognition process which comprises two main stages that are signal processing and pattern classification. The sound recorder, webcam &MyRio setup is placed near incubator of infant in NICU.

The system receives the raw cry signal as the input through the microphone. In the signal processing step the raw cry signal is pre-processed (using Labview program), normalized and filtered. The smoothed signal is analyzed by suitable feature extraction called multi tone information technique. One of the important issues in cry analysis is length of the segmented cry samples. Therefore, cry segments of fixed length are extracted from the cry records although the cry utterances are not equal in length. This cry is further processed using FFT module for more information on frequency, pitch, samples other details. The same details of voice signal, FFT signal graph, & multi tone information with vision display of webcam will be done on programming window as a result too. This process is made to run continuously with while loop & a control unit for stopping process. This processed signal is further given to logical comparison program for the classification of the pain level assessment. The resultant infant will be displayed in terms of string message with the cry signal graph, tone information, webcam data display.

These results can be viewed in 2 modes such as work station/ pc display, and another as Smartphone/ tablet display. In work station display, execution of the program will show the entire result on the same pc screen. For the Smartphone display the entire MY Rio setup is connected to the same WiFi network, using this with help of Labview data dashboard we can view the results on Smartphone (graph, message notification, & webcam data display), so it will be helpful for nurses or caring unit to monitor the infants easily.

Figure 6. Labview configuration module



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Advantage of this program is that the care taker/ nurse need not be 24/7 near the babies this can be monitored by remotely via mobile/tabs which should be connected on the same wifi network. The baby monitoring system can be viewed in the workstation laptop which intern controls MYRIO and on doctors' smart phone as shown below.

Figure 7. Workstation displaying the reason for cry, the graph charts of feature extraction and baby image at that instant

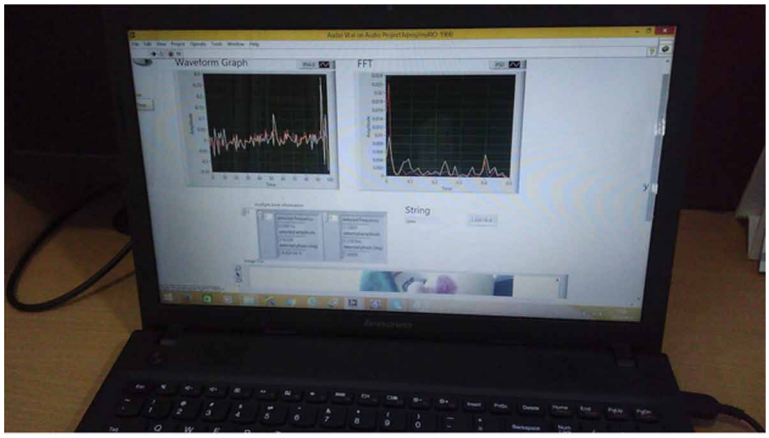


Figure 8. Continuous baby monitoring visualization in smart phone/ doctors phone

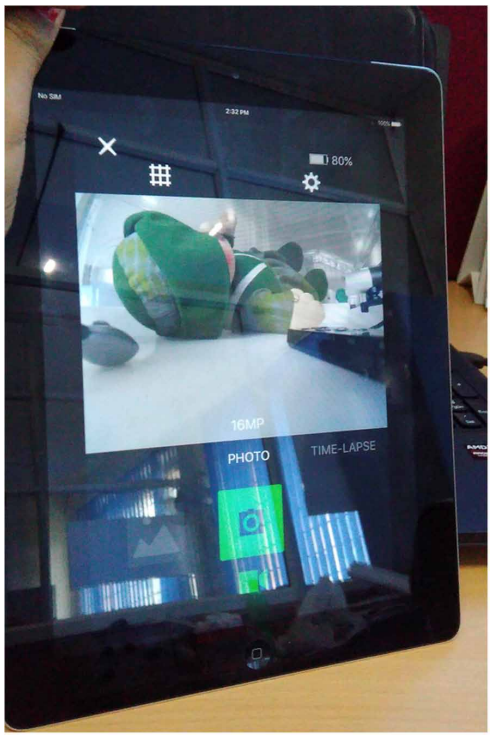
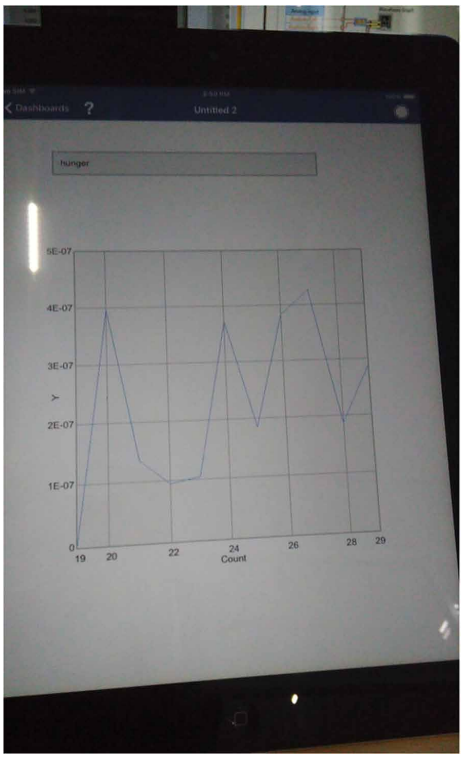


Figure 9. Reason of cry and feature extraction graph displayed on doctor smart phone



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The display of message and classification data can be done in 2 parts one on the same screen where program is running & other one on the mobile/tabs which will be connected to the program variables of Labview via the same wifi network. This web services can be made available for concern person to access the data. The Labview data dashboard is used for displaying the data remotely. The webcam data is displayed using YICAM app/ ip cam.

DISCUSSION

The proposed study simulates the scenario of the typical NICU environment. The experimental work was conducted to emphasize the type of cry signal and conducted to video snap shot of the infant kept inside the incubator in the NICU. The voice recorder with a baby phantom kept inside the incubator and the corresponding type of cry was recognized. The real-time myRio processor with a Labview environment provides the graphical interface. Finally, the LCD tablet/mobile which was interfaced with the Laptop shows the display of the phantom (infant) as well as the voice record stating the type of cry. Such arrangements help the staff nurse/duty clinician to attend the baby immediately to cater their needs. Attempts are under process for executing the same process in the real-time scenario at NICU. The performance of the real-time processor is currently ongoing in terms of testing its overall efficiency, processor speed, memory capabilities etc.

CONCLUSION

In a neonatal intensive care unit (NICU), there is a huge need to monitor the physiological conditions of the premature and full term infants. In such environment, the cry signal conveys significant information in terms of the distress of the infant and thereby needs of the infants can be well attended. Establishment of the portable device is thus essential for such scenario. This work suggested a primitive experimental study on developing portable electronic medical device for infant cry recognition. The cry signals were recorded by placing the voice recorder in the incubator/warmer and three variants, hunger wet diaper, and pain were considered for the proposed study. All the recordings were qualitatively assessed by the clinician and confirm its suitability for the experimental study. The real time My Rio processor along with the labview software performed the operation of pre- processing and feature extraction and a threshold based classification was employed which sends the voice based alarm to the tablet. The continuous monitoring of such facility in the centralized nursing station in NICU provides the behavioral status of the infant and thus immediate attention for the clinical needs can be executed without any further delay.

ACKNOWLEDGMENT

The authors would like to thank Vision Group of Science and Technology, Government of Karnataka, Bangalore, Karnataka, India for their financial support to carry out this research project (SMYR/2014-2015/GRD No. 294). The authors would like to acknowledge Department of Pediatrics, M S Ramaiah Medical College & Hospital, Bangalore, India for the assistance and help during recording the cry rhythm.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Bard, D., Runefors, P., & Arnbjornsson, E. (2008). Newborn pain cry analysis based on pitch frequency tracking. University Hospital Lund, Lund, Sweden.

Chang, C.-Y., Chang, C.-W., Kathiravan, S., Lin, C., & Chen, S.-T. (2016). DAG-SVM based infant cry classification system using sequential forward floating feature selection. *Multidimensional Systems and Signal Processing*, April.

Messaoud, A., & Chakib Tadj (2011, May 8-11). Analysis of Acoustic Features of Infant Cry for Classification Purpose. *Proceedings of the 24th Canadian Conference On Electrical And Computer Engineering CCECE*, Niagara Falls, ON, Canada. 10.1109/CCECE.2011.6030415

Rautava, L., Lempinen, A., Ojala, S., Parkkola, R., Rikalainen, H., Lapinleimu, H., ... Lehtonen, L. (2007). Acoustic quality of cry in very-low-birth-weight infants at the age of 1 1/2 years. *Early Human Development*, *83*(1), 5–12. doi:10.1016/j.earlhumdev.2006.03.004 PMID:16650947

Reyes-Galaviz, O. F., & Reyes-Garcia, C. A. (2004, September 20-22). A System for the Processing of Infant Cry to Recognize Pathologies in Recently Born Babies with Neural Networks. *Proceedings of the 9th Conference Speech and Computer SPECOM'04, St. Petersburg*, Russia.

Sandhya, S. Jagtap, Premanand K Kadbe, (2016). System proposed for be acquainted with newborn cry emotion using linear frequency Cepstral coefficient. *Proc. of International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*.

Sriraam, N., & Tejaswini, S. (2014). Infant Cry Analysis and Pain Scale Management: A Pilot Study. *International Journal of Biomedical and Clinical Engineering*, *3*(1), 42–51. doi:10.4018/ijbce.2014010104

Talauliker, D.U.S., & Shenvi, N. (2015). Analysis of Cry in New Born Infants. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(3).

Bhagatpatil, V.V., & Sardar, V.M. (2014). An Automatic Infant's Cry Detection Using Linear Frequency Cepstrum Coefficients (LFCC). *International Journal of Scientific & Engineering Research,* 5(12).

Warlaumont, A.S., Oller, D.K., & Buder, E.H. (2010, April). Data-driven automated acoustic analysis of human infant vocalizations using neural network tools. *J. Acoust. Soc. Am.*, 127(4), 2563-2577

Zabidi, A., Mansor, W., Lee, Y.K., MohdYassin, A.I., & Sahak, R. (2010, October 16-18). Particle Swarm Optimisation of Mel-frequency Cepstral Coefficients Computation for the Classification of Asphyxiated Infant Cry. *Proceedings of the 3rd International Conference on Biomedical Engineering and Informatics*, Yantai. 10.1109/BMEI.2010.5639674

This research was previously published in the International Journal of Biomedical and Clinical Engineering (IJBCE), 5(2); edited by Natarajan Sriraam; pages 53-63; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 16
Efficient String Matching Algorithm for Searching Large DNA and Binary Texts

Abdulrakeeb M. Al-Ssulami
King Saud University, Saudi Arabia

Hassan Mathkour
King Saud University, Saudi Arabia

Mohammed Amer Arafah
King Saud University, Saudi Arabia

ABSTRACT

The exact string matching is essential in application areas such as Bioinformatics and Intrusion Detection Systems. Speeding-up the string matching algorithm will therefore result in accelerating the searching process in DNA and binary data. Previously, there are two types of fast algorithms exist, bit-parallel based algorithms and hashing algorithms. The bit-parallel based are efficient when dealing with patterns of short lengths, less than 64, but slow on long patterns. On the other hand, hashing algorithms have optimal sublinear average case on large alphabets and long patterns, but the efficiency not so good on small alphabet such as DNA and binary texts. In this paper, the authors present hybrid algorithm to overcome the shortcomings of those previous algorithms. The proposed algorithm is based on q-gram hashing with guaranteeing the maximal shift in advance. Experimental results on random and complete human genome confirm that the proposed algorithm is efficient on various pattern lengths and small alphabet.

1. INTRODUCTION

The problem of finding the occurrences of a predefined pattern in big data or very long sequences is classical problem in computer science and it has been studied intensively by researchers due to the growing of data which are produced day-by-day. There are many applications that depend on the pat-

DOI: 10.4018/978-1-7998-1204-3.ch016

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tern matching, e.g., biological sequences which are produced every day by high throughput sequencing technologies (Morozova & Marra, 2008), natural language processing, or retrieving a specific data from very large databases. Bioinformatics is one of the fields where the pattern matching problem is used widely to search for specific pattern exactly or approximately (Barton, Iliopoulos, & Pissis, 2014; Simone Faro & Lecroq, 2009a; S. Faro & Lecroq, 2012; Kalsi, Peltola, & Tarhio, 2008) where the DNA is considered as a string and the four symbols *A*, *C*, *G*, and *T*, represent the four nucleotides Adenine, Cytosine, Guanine, and Thymine, respectively. Computer networks is the other area where this problem is applied. The binary string matching used in the intrusion detection systems to detect malwares within the heavy load traffic and searching IP addresses in routers swiftly (Chu, Huang, Tsai, & Hsieh, 2008; Liu, Huang, Chen, & Kao, 2004; Tuck, Sherwood, Calder, & Varghese, 2004).

The importance of speeding-up the searching process stems from that there is a need to find the occurrences of a particular pattern and whether the pattern exist or not in terabytes of data. Searching such amount of data offline by uploading the data to memory is not possible due to the memory limitations. In computer networks, Antiviruses are installed on routers or personal computers for detecting malwares so speeding-up the inspecting process increases the performance of the systems.

The problem is formulated as follows. Given a string T of length n and a pattern P of length m, the problem is to find the number of repetitions of P in T. In practice $m \ll n$. Let Σ denotes the set of symbols, then Σ^* denotes the language from which the pattern and the text are drawn and we write $P,T \in \Sigma^*$.

Two types of alphabets are considered in this paper, the DNA alphabet, $\Sigma=\{A,C,G,T\}$, and binary alphabet, $\Sigma=\{0,1\}$. The text and the pattern are represented by arrays of one dimension and the two arrays are indexed starting from 0. We say that the pattern P occurs in the text T, if $P[i]=T[j-m+i+1]$ where $0 \leq i < m$ and $\exists j$ such that $m-1 \leq j < n$.

2. LITERATURE REVIEW

There are so many exact string matching algorithms which are varying between character-based comparison(Al-Ssulami, 2015; Boyer & Moore, 1977; Deusdado & Carvalho, 2009; Franek, Jennings, & Smyth, 2007; Horspool, 1980; James H. Morris & Pratt, 1970; Knuth, James H. Morris, & Pratt, 1977; Lecroq, 2007; Sunday, 1990; Tarhio & Peltola, 1997), automata-based (Allauzen, Crochemore, & Raffinot, 1999; Allauzen & Raffinot, 2000; Blumer et al., 1985; Fan, Yao, & Ma, 2009; He, Fang, & Sui, 2005; Simon, 1994), and bit-parallel based (Baeza-Yates & Gonnet, 1992; Branislav, Durian, Holub, Peltola, & Tarhio, 2010; Cantone, Faro, & Giaquinta, 2010; Chen, Huang, & Lee, 2014; Simone Faro & Lecroq, 2009b; Mohanty & Tragoudas, 2014; Navarro & Raffinot, 2000; Peltola & Tarhio, 2003, 2014; Wu & Manber, 1992). Majority of character-based algorithms depend on the previously published ideas in (Boyer & Moore, 1977; James H. Morris & Pratt, 1970; Knuth et al., 1977) (See Simone Faro & Lecroq, 2013). The automata-based string matching depends on the idea of the finite automata. The algorithms of this type build the minimal deterministic finite automat then the text is parsed with the automata in time complexity of $O(n)$ in case of using an array of size $m|\Sigma|$; otherwise, the search phase costs $O(n\log|\Sigma|)$. The deterministic finite automata DFA is built for the pattern P of length m in time and space complexity of $O(m|\Sigma|)$ if an array is used (See (Simone Faro & Lecroq, 2013)). The bit-parallel based algorithms take the advantage of parallelism the bits inside the computer word. This technique

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

reduces the number of operations by a factor of ω, where ω is the size of computer word, that is, ω comparisons can be performed in a single bitwise comparison. This technique is suitable for simulation of the Non-deterministic finite automata (Simone Faro & Lecroq, 2013).

Recently, A bit-parallel algorithms for binary and DNA random texts is presented by (Peltola & Tarhio, 2014). The authors generalized the FSBNDM (Forward Simplified Backward Non-deterministic DWAG Matching) algorithm to inspect more than one character beyond the current text window. The FSBNDM algorithm by (Simone Faro & Lecroq, 2009b) reads two characters $T[j]$ and $T[j+1]$, where j is the last position in the current text window. Therefore, there are three situations:

1. $T[j]=P[m-1]$, then the reading continues leftwards;
2. $T[j]T[j+1]$ is a factor of P, reading continues leftwards;
3. $T[j]T[j+1]$ is not a factor of P, shifting the text window right by $m \leq 64$ positions (larger shift than SBNDM2 by 1).

(Peltola & Tarhio, 2014) developed *FSB*(q,f) algorithm to do the tests in cases 1 and 2 in one step. Based on the observation that when the q-gram is increased, the efficiency of SBNDMq (Branislav et al., 2010) is also increased. *FSB*(q,f) has been designed to use larger q-gram than 2 and handling f lookahead characters instead of 1, where $0 \leq f < q$.

Another algorithm called Greedy-SBNDM2 (GSB2) is also presented by Peltola and Tarhio (Peltola & Tarhio, 2014) to use a greedy skip loop or bad character heuristics. As a result, the authors proved experimentally that the algorithms *FSB*(q,f) and GSB2 are faster on binary and DNA data. Recently, the SSM algorithm is presented by Al-Ssulami (Al-Ssulami, 2015) to be faster in practice for texts of large alphabets.

The main drawback of *FSB*(q,f) and GSB2 algorithms is that their efficiency is restricted to the size of computer word 64, while SSM and HASHq (Lecroq, 2007) are efficient for large alphabet and long patterns but suffer from small shift in case of binary and DNA data. So in this paper, we present an efficient algorithm to deal with short, long and very long patterns on both binary and DNA texts. The new algorithm, namely HSSMq, consists of determining the q-gram pivot position, initializing the hash table and computing the possible shifts, maximizing the safe shift, and searching by matching the hash values. If the hash value of the q-gram pivot equals the hash value of the corresponding text q-gram, the verification step of matching starts from the rightmost end of the pattern towards the leftmost end.

The experimental results on binary and DNA random texts exhibit that the new algorithm is faster than the recent algorithms on short, long and very long pattern lengths.

The rest of this paper is organized as follows. Section 3 introduces in details HASHq with some figures to show its behavior on different alphabet sizes then the SSM algorithm is introduced with a figure shows how much it was improved over Horspool that we consider as HASH1. After that the pseudo-code of the new algorithm is sketched with experimental results showing how much is improved over HASHq by computing the average shift length, Section 4 shows the experimental results, and Section 5 provides our conclusion and future work.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3. THE ALGORITHMS

In more details, we introduce the well-known and fast algorithm HASHq (Lecroq, 2007) and the recent one (Al-Ssulami, 2015) because they are directly related to the new algorithm. At the end of this section, we present the new algorithm with pseudo-codes and an experiment showing the improvement on the shift length.

3.1. HASHq Algorithm

Let $\hbar_k \in \{0..255\}$. Then ∀ *q*-gram in the pattern *P*, the algorithm computes

$$\hbar_k = \left(\sum_{i=0}^{q-1} 2^i P[k-i]\right) mod 256,$$

where $q\text{-}1 \le k \le m\text{-}2$ and *m* is the length of *P*. The amount of shift is computed and stored in *Shift*[] array as:

$$Shift[\hbar_k] = \begin{cases} m-k-1, & k = \max\{q-1 \le l < m-1 : \hbar_l = \hbar_k\} \\ m-q+1, & if\ \ no\ such\ that\ \hbar_l = \hbar_k \end{cases} \quad (1)$$

The algorithm stores the amount of shift for rightmost *q*-gram in variable $sh_1 = Shift[\hbar_{m-1}]$ and updates $Shift[\hbar_{m-1}] = 0$.

Let us define the current text window as a portion of the text *T* of length *m* where a pointer *j* refers to the rightmost symbol in the window and $m\text{-}1 \le j < n$.

This algorithm uses fast skip loop. It compares the hash value of the rightmost *q*-gram of the current text window with 0. If they are not equal, the current text window slides to right by $Shift[\hbar_j]$; otherwise, it slides by sh_1.

A problem with this algorithm arises when the alphabet is small which means that the value $m\text{-}k\text{-}1$ in Equation (1) is small.

Following is a computation for the expected shift value with different alphabet sizes. In (Baeza-Yates, 1989), the average expected shift length of Horspool algorithm is computed as follows:

$$Shift_{(Horspool)} = |\Sigma|\left(1 - \left(1 - 1/|\Sigma|\right)^m\right) \quad (2)$$

Similarly, we compute the expected shift length of HASHq algorithm by the following equation considering that the hashing function is free of collisions:

$$Shift_{(HASHq)} = |\Sigma|^q \left(1 - \left(1 - 1/|\Sigma|^q\right)^{m-q+1}\right) \quad (3)$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The average expected shift lengths for different alphabet sizes, pattern lengths, and q-grams are shown in Figures 1-3. It turns out that the average shift length is very small when $|\Sigma| \leq 5$.

The figures confirm that HASHq is indeed fast for large alphabet but it suffers from small shift in case of binary and DNA alphabets (small alphabets).

3.2. SSM Algorithm

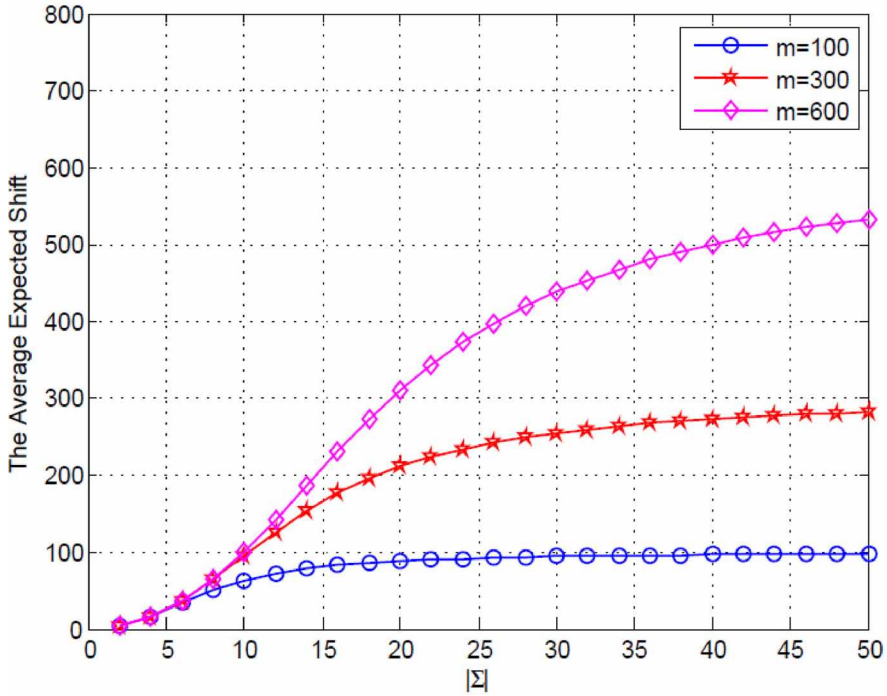
We found the recent algorithm SSM is suitable to be adapted for small alphabets. It seeks for a pivot position ν in the pattern P such that $P[\nu]$ has maximal shift γ. Equation (4)-(6) describe the pivot ν with the maximal shift γ:

$$\gamma_1 = \begin{cases} \max\left\{\left(\nu - i\right) : 0 \leq i < \nu < m \, and \, P\left[i\right] = P\left[\nu\right]\right\} \\ 1, \quad otherwise \end{cases} \tag{4}$$

$$\gamma_2 = \begin{cases} \max\left\{\left(\nu + 1\right) : where \, P\left[\nu\right] is \, unique \, in \, P\right\} \\ 1, \quad otherwise \end{cases} \tag{5}$$

$$\gamma = \max\{\gamma_1, \gamma_2\} \tag{6}$$

Figure 1. The average expected shift length of HASHq when q=2



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The average expected shift length of HASHq when q=3

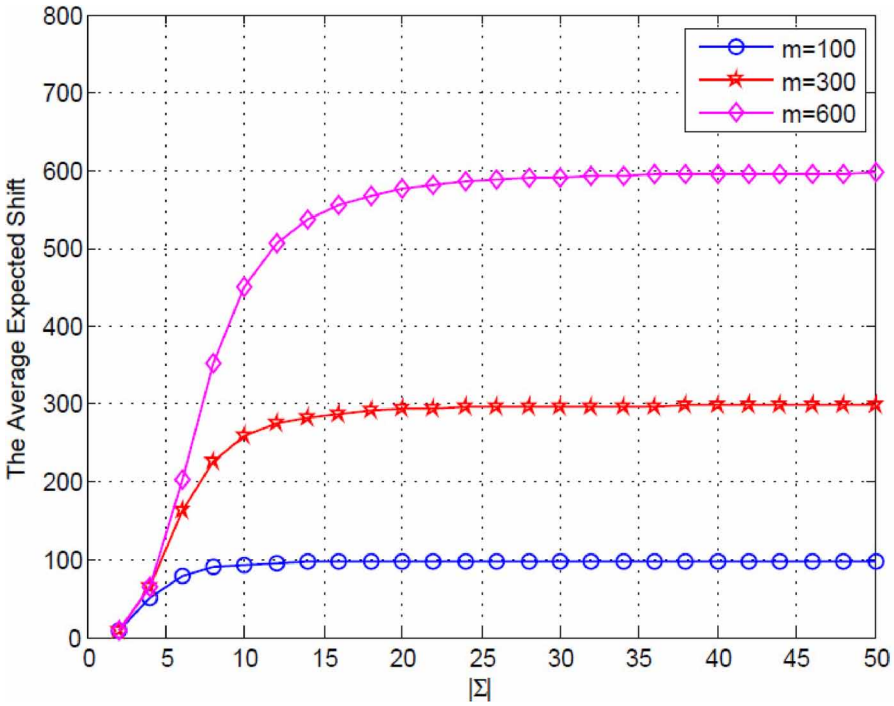
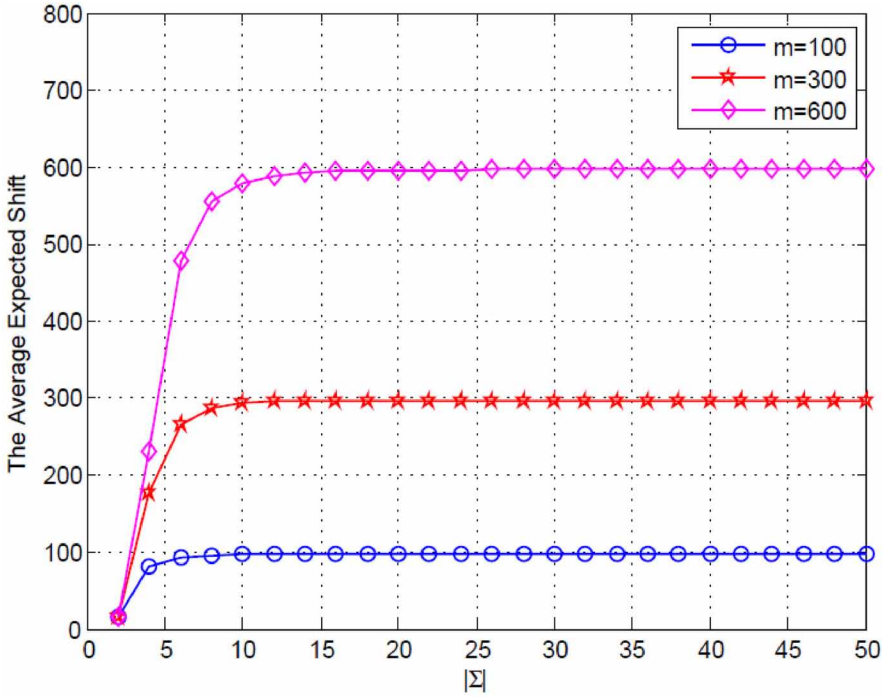


Figure 3. The expected shift length of HASHq when q=4



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

SSM algorithm increases the amount of shift by scanning the pattern from left to right and vice versa which is suitable for long runs. This shift is stored in array *Shift*[]. Finally, a maximization step is made to compare the γ with each value in *Shift*[]and within each step the maximum shift is stored.

This algorithm uses a hash array *Horspool*[]where its entry is a single character, like Horspool. The skip loop condition is a comparison between $P[\nu]$ and $T[j+\nu-m+1]$, where j is the rightmost position of the current text window. If a match occurs, then the matching process is performed from the rightmost end of the pattern towards the leftmost end. This process ends either with a match or a mismatch at any position k and then the window is slided right by *Shift*[]. In case a mismatch occurs at ν, the shift is by *Horspool*[$T[j]$].

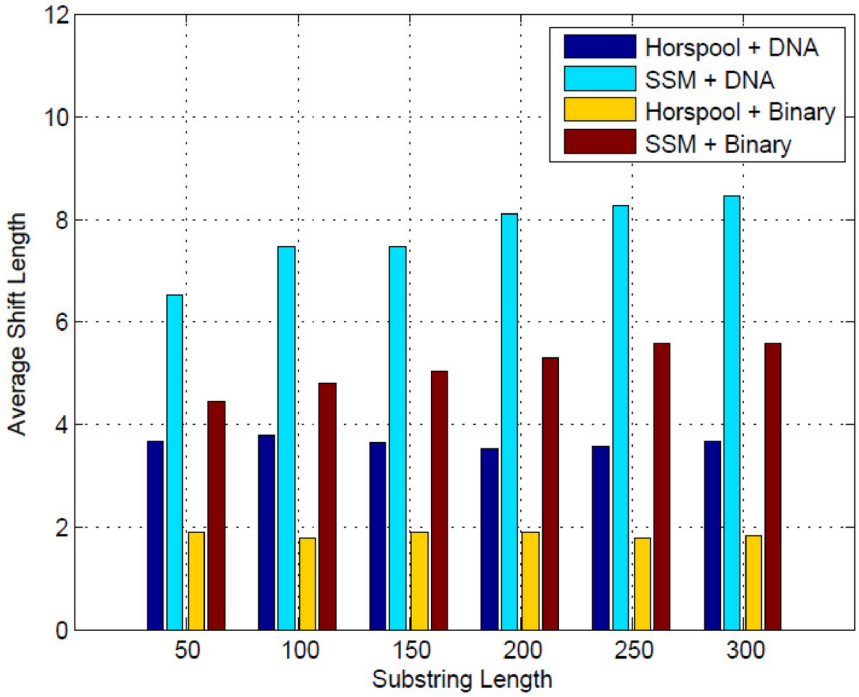
Experimentally, we measured the average shift length of Horspool and SSM algorithms on lengths 50-300 step 50. The results are shown in Figure 4 and confirm that the SSM added improvement over Horspool which can be transferred to HASHq.

The results show that SSM algorithm indeed has longer average shift length than Horspool. From this point, we found that it is possible to extend the idea of SSM to hashing more than a single character and getting better results on both dataset types. So, the new proposed algorithm (HSSMq) is a combination of hashing like HASHq and guaranteeing the maximal shift in advance like SSM algorithm.

3.3. The Proposed Algorithm

As in Figures 1-3, HASHq has longer shift on large alphabets and long patterns but it is not so for small alphabets. In the meanwhile, SSM more efficient than Horspool on small alphabets (See Figure 4) which

Figure 4. The average shift length over 100 runs



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

we consider as a special case of HASHq when q=1. Based on this conclusion, we design the HSSMq to integrates the idea of SSM with the q-gram hashing with some modifications.

The new algorithm consists of two steps, the preprocessing (Algorithms 1- 2) and searching phase (Algorithm 3). The preprocessing phase is represented by two subphases, the initialization and the preprocessing to compute the possible shifts.

Formally, let $Hash_Array_q[]$ denotes the hash array of $\sigma=2^{16}$ entries. Notice that HASHq has smaller size, 2^8. Increasing the size of hash array reduces the hash collisions and hence increasing the shift length. The hash array $Hash_Array_q[]$ is initialized with the maximum shift length $\chi=m-q+1$, (Lines 6-7, Algorithm 1).

Let us define the q-gram pivot ν as the q-gram of the pattern where its rightmost symbol is positioned at ν and gives a maximal shift length γ. So we can rewrite Equation (4)-(6) as:

$$\gamma_1 = \begin{cases} \max\left\{\left(\nu - i\right) : q-1 \le i < \nu < m \, and \, \hbar_i = \hbar_\nu\right\} \\ 1, \quad otherwise \end{cases} \tag{7}$$

$\hbar_i$ is the hash value for the q-gram at position i of the pattern which is computed by Lines 10-14 of Algorithm 1. The value of (ν-i) is represented by Line 18 of Algorithm 1:

$$\gamma_2 = \begin{cases} \max\left\{\left(\nu - q + 2\right) : \quad where \hbar_\nu \, is \, unique \, in \, P\right\} \\ 1, \quad otherwise \end{cases} \tag{8}$$

Equation (8) is represented by Line 16 of Algorithm 1:

$$\gamma = \max\{\gamma_1, \gamma_2\} \tag{9}$$

Line 20-21 of Algorithm 1 represents Equation (9). The pivot position ν is then chosen accordingly.

Equations (7)-(9) determine the pivot ν. The next step is reinitializing $Hash_Array_q[]$ according to Equation (1) of HASHq, but with the new pivot ν as follows:

$$Hash_Array_q\left[\hbar_k\right] = \begin{cases} \nu - k, & where \;\; k = \max\left\{q-1 \le l \left\langle \nu \right| \hbar_l = \hbar_k\right\} \\ \nu - q + 2, & if \;\; no \, such \;\; that \;\; \hbar_l = \hbar_k \end{cases} \tag{10}$$

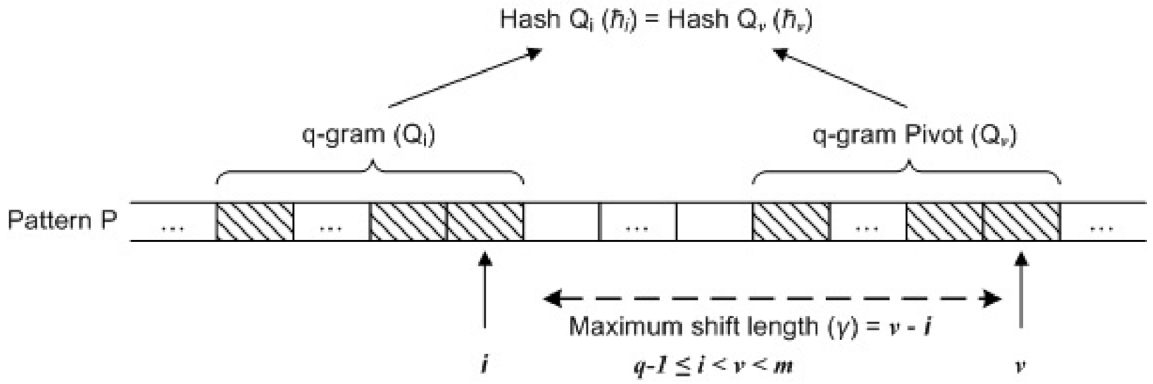
Equation (10) corresponds to Lines 22-30 of Algorithm 1.

Algorithm 2 is the preprocessing of the SSM algorithm. It consists of three steps, computing the shift length from left towards right side (Lines 2-7), updating the shift length from right to left (Lines 8-15), and the maximization step which updates the array $Dist[]$ so that $Dist[i]=\max\{\gamma, Dist[i]\}$ and $0 \le i < m$ as stated in lines 16-19.

Graphically, Figure 5 depicts the pivot q-gram (Q_ν) at position ν which has a maximum distance (γ) of safe shift so that its left q-gram (Q_i) that has the same hash value $\left(\hbar_\nu\right)$ is positioned at i in the pattern where there is no q-gram between i and ν has the same hash value $\hbar_\nu$.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The q-gram pivot with maximum safe shift



The searching phase is sketched in Algorithm 3. Let *j* is the position of the rightmost symbol in the current text window and ranges over text *T* such that $m\text{-}1 \le j < n$. Thus, $\hat{j} = j + \nu - m + 1$ is the text window position that corresponds to ν. $\hbar_{\nu}$ and $\hbar_{\hat{j}}$ are the hash values which are computed according to Line 9 and Line 12, respectively.

Within each jump for *j*, $\hbar_{\nu}$ is compared with $\hbar_{\hat{j}}$. If they are equal, a verification step is triggered (Lines 16-19, Algorithm 3, then *j* is increased by $Dist[i] \le \gamma$ (Line 24)). If they are differ, *j* jumps $Hash_Array_q\left[\hbar_{\hat{j}}\right]$ positions (Line 14).

In recap, if the current text window position is $m\text{-}1 \le j < n$, then the next position $\bar{j}$ is computed by Equation (11):

$$\bar{j} = \begin{cases} Dist[i] + j & if\ \hbar_{\hat{j}} = \hbar_{\nu} \\ Hash_Array_q\left[\hbar_{\hat{j}}\right] + j & otherwise \end{cases} \tag{11}$$

Figure 6 illustrates briefly the layout of the searching process and shows how the different variables are positioned in both pattern and text. As shown from the figure, at position $\hat{j}$, the hash values are not equal $\left(\hbar_{\nu} \neq \hbar_{\hat{j}}\right)$, so a mismatch occurs at that position resulting in sliding the text window to right skipping a set of text symbols without comparisons. In the next attempt, a match, $\left(\hbar_{\nu} = \hbar_{\bar{j}}\right)$, of the hash values occurs at position $\bar{j}$. Thus, a verification step must be performed to make sure whether an exact match occurs or not. This step is performed by matching the characters one by one.

By analyzing the time complexity of the algorithm, it is obvious that the preprocessing time complexity is $O(mq)$, where *m* is the length of pattern and *q* is the size of q-gram and the searching phase takes time complexity of $O(n/m)$ for the average and optimal cases while the worst case is $O(nm)$ and that is the same in general for HASHq, SSM, bit-parallel algorithms.

We measured the average shift length of the two algorithms HASHq and HSSMq on DNA and binary random texts of size 200 MB and 100 runs. The results in Figures 7-8 show that HSSMq has longer shift on both texts.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The layout of the searching phase of the HSSMq algorithm

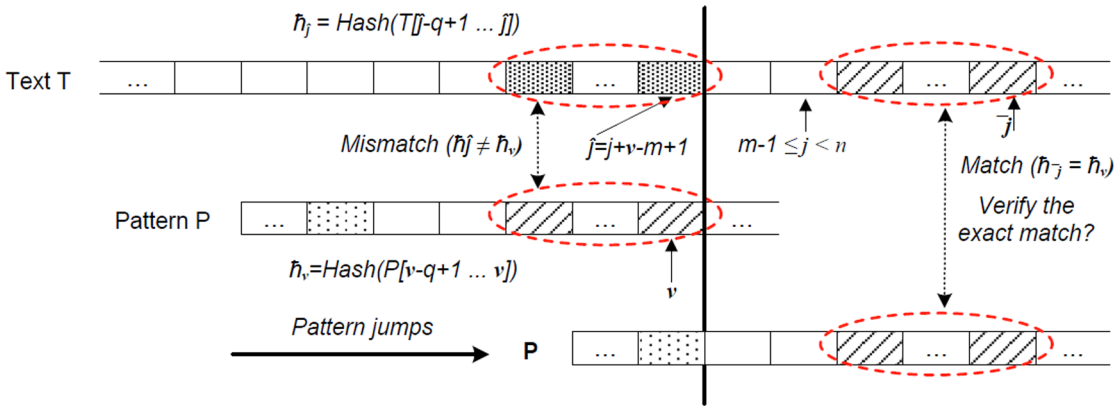
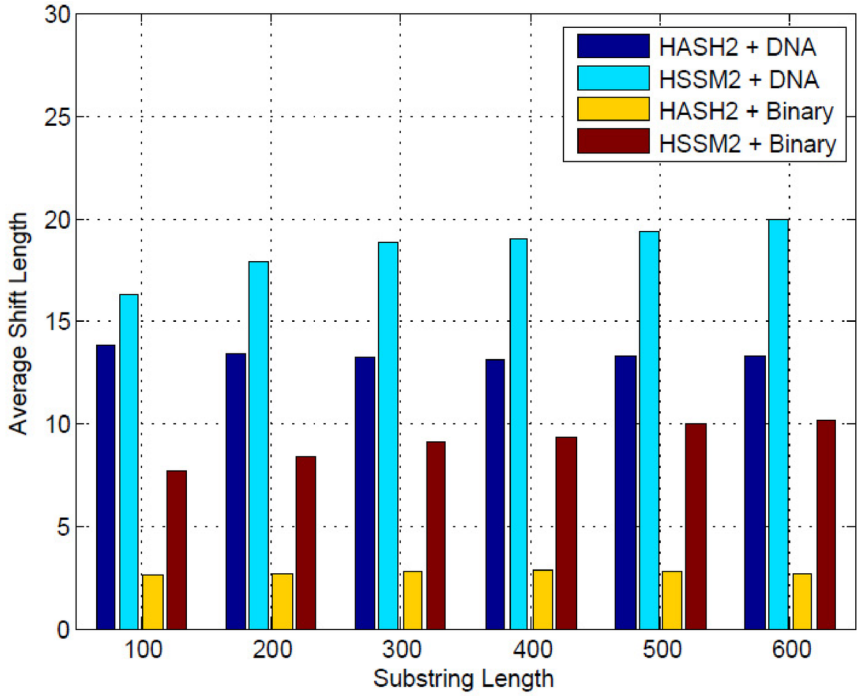
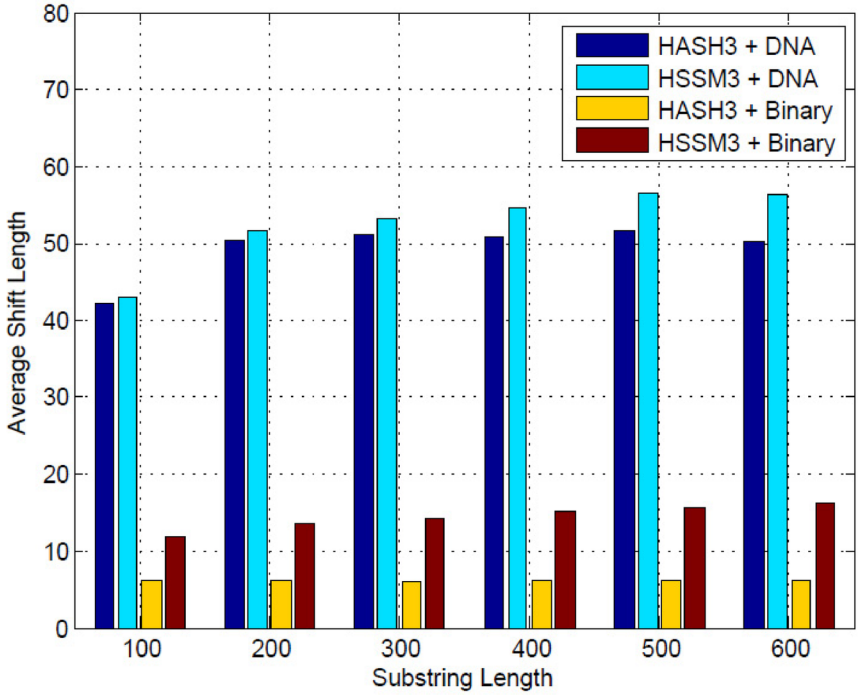


Figure 7. Average shift length when q = 2



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Average shift length when q = 3



Algorithm 1. Initializing the hash table and determining the pivot v

Input: P, m and q.
Output: The pivot ν, $Hash_Array_q[]$, and γ.

```
1  begin
2      σ ← 2^16  // size of Hash_Array_q[] array.
3      χ ← m − q + 1,  μ_1 ← m − 1,  ν ← 0
4      if (m < q) then
5          Output "Input Error"
6      for  i ← 0, 1, ..., σ − 1 do
7          Hash_Array_q[i] ← χ
8      γ ← 1,  Λ ← 0
9      for i ← q − 1, q, ..., μ_1 do
10         ħ_i ← P[i − q + 1]
11         ħ_i ← (char)((ħ_i ≪ 1) + P[i − q + 2])
12         ...
13         ħ_i ← (char)((ħ_i ≪ 1) + P[i − 1])
14         ħ_i ← (char)((ħ_i ≪ 1) + P[i])
15         if  (Hash_Array_q[ħ_i] = χ) then
16             Λ ← i − q + 2
17         else
18             Λ ← i − (μ_1 − Hash_Array_q[ħ_i])
19         Hash_Array_q[ħ_i] ← μ_1 − i
20         if (Λ ≥ γ) then
21             γ ← Λ,  ν ← i
22     for i ← 0, 1, ..., σ − 1 do
23         Hash_Array_q[i] ← ν − q + 2
24     for i ← q − 1, q, ... ν − 1 do
25         ħ_i ← P[i − q + 1]
26         ħ_i ← (char)((ħ_i ≪ 1) + P[i − q + 2])
27         ...
28         ħ_i ← (char)((ħ_i ≪ 1) + P[i − 1])
29         ħ_i ← (char)((ħ_i ≪ 1) + P[i])
30         Hash_Array_q[ħ_i] ← ν − i
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 2. The preprocessing phase of HSSM$_q$ algorithm

```
Input: P, m.
Output: The pivot ν, Hash_Array_q[], and array Dist[].
1  begin
2      Dist[0] ← 1, ρ ← 1
3      for i ← 1, 2, ..., m − 1 do
4          if P[i] ≠ P[i − 1] then
5              Dist[i] ← 1, ρ ← 1
6          else
7              ρ ← ρ + 1, Dist[i] ← ρ
8      ρ ← 1
9      for i ← m − 1 down to 1 do
10         if (P[i] ≠ P[i − 1]) then
11             if (ρ ≥ Dist[i − 1]) then
12                 Dist[i − 1] ← ρ + 1
13             ρ ← 1
14         else
15             ρ ← ρ + 1
16     (Hash_Array_q[], ν, γ) ← Algorithm 1
17     for i ← 0, 1, ..., m − 1 do
18         if (Dist[i] < γ) then
19             Dist[i] ← γ
```

Algorithm 3. The searching phase of HSSM$_q$ algorithm

```
Input: Pattern P of length m and Text T of length n.
Output: Count, the number of occurrences.
1  begin
2      ( ν, Hash_Array_q[], Dist[] ) ← Algorithm 2
3      j ← m − 1
4      j_0 ← ν − m + 1
5      j_1 ← j_0 − 1
6      ...
7      j_qMinus2 ← j_0 − q + 2
8      j_qMinus1 ← j_0 − q + 1
9      ħ_ν ← (char)(((((P[ν − q + 1] ≪ 1) + ...) + P[ν − 1]) ≪ 1) + P[ν])
10     T ← T + P
11     while (j < n) do
12         ħ_j ← (char)(((((T[j + j_qMinus1] ≪ 1) + ...) + T[j + j_1]) ≪
           1) + T[j + j_0])
13         while (ħ_j ≠ ħ_ν) do
14             j ← j + Hash_Array_q[ħ_j]
15             ħ_j ← (char)(((((T[j + j_qMinus1] ≪ 1)...) + T[j + j_1]) ≪
               1) + T[j + j_0])
16         i ← m − 1
17         Pos ← 1 − m + j
18         while ((i ≥ 0) and (P[i] = T[i + Pos])) do
19             i ← i − 1
20         if (i < 0) then
21             if (j < n) then
22                 count ← count + 1
23             i ← i + 1
24         j ← j + Dist[i]
25     return Count
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. EXPERIMENTAL RESULTS

We conducted the experiments on a PC Intel® Core™ i7 x64-based CPU running at 3.40 GHz and 8 GB RAM and 64-bit Windows 8.1 Pro. All the algorithms are implemented in C# programming language. The implementations of the other algorithms which are used in this study are obtained from the web site by Faro and Lecroq[1]. For all bit-parallel algorithms in this study, we used a bit vector of 64 bits.

We generated a random dataset of size 200 MB where the symbols of each alphabet are statically independent of each other and equiprobable. The symbols of the binary dataset are 0 and 1 and the DNA dataset contains the symbols A, C, G, and T. For each dataset, the program selects 200 patterns randomly from the given data and computes the average running time in thousands of a second. We divided the length of patterns into three groups short (10-80), long (100-800), and very long patterns (1,000-5,000).

The best running times are underlined and boldfaced in all tables. In the experiments, we chose to compare with the related and recent algorithms SSM, SBNDMq, FSBNDM, FSB(q,f), GSB2, and HASHq.

Table 1 shows the running times on the binary dataset for short patterns. The results show that HSSM4 reported the best running time on length 10 and HSSM6 has the best times on lengths 20-80 followed by HASH4 on length 10 and HASH6 on lengths 20-80. Among all the bit-parallel algorithms, FSB(6,2) is the best on length 10, SBNDM4 is better on length 20, FSB(8,3) on length 30, FSB(8,2) reported better running times on lengths 40-60 followed by FSB(8,1) on lengths 70-80. We scaled the running times in Figure 9 to show the difference between these algorithms.

We can say that HSSMq algorithm is a hybrid of SSM and HASHq algorithms. HSSM2 is a hybrid of SSM and HASH2 and so on. Regarding Table 1, it turns out that the running time of SSM is declined as follows. The running time is reduced by 14-42% compared to HSSM2, HSSM3 decreased the running time by 39-65%, with HSSM4 the running time is reduced by 48-78%, and 41-89% with HSSM6.

In Table 1, we also see that HSSM2 reduced the running time of HASH2 by 6-55%, the running time of HSSM3 versus HASH3 is decreased by 2-45%, running time of HSSM4 versus HASH4 is decreased 5-29%, and the running time of HSSM6 versus HASH6 is decreased by 18-21%. Likewise, the running time of HSSM6 versus FSB (8,2) is decreased by 18-33%.

On long patterns and binary dataset, it turns out from Table 2 that HSSM6 is the best overall followed by HASH6 and FSB (8,1). The difference between them is shown in Figure 10.

Likewise, Table 3 presents the running times on very long patterns and binary dataset and the difference in the efficiency between HSSM6, HASH6, and FSB(8,1) is shown in Figure 11.

Considering the DNA text and short patterns, HSSM3 is the best on lengths 10-20, HSSM4 won on lengths 30-60, and HASH6 on lengths 70-80 (see Table 4). Among the remaining algorithms excluding HSSMq class, FSB(3,1) won the second place on length 10, SBNDM4 on lengths 20-70, and HASH4 on length 80. The best running times of these algorithms have been scaled using $\log_2()$function to see the difference in Figure 12.

On DNA and long patterns, it is obvious from Table 5 that HSSM6 is the winner algorithm followed by HASH6. Among those bit-parallel algorithms, SBNDM4 is the best. The difference is clearly shown in Figure 13. The efficiency of these three algorithms is also continued with very long patterns as it is shown in Table 6 and it is illustrated by Figure 14.

It seems that on DNA text as the length of pattern is increased, the difference in time between HSSM6 and SBNDM4 is also increased and the difference between HASH6 and HSSM6 tends to be closer. Also we can see that on lengths longer than 64, the efficiency of bit-parallel algorithms tends to be stable or somewhat fixed because that the efficiency is restricted to the computer word length, 64.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. The average running times on random binary dataset of size 200 MB and short

	10	20	30	40	50	60	70	80
SBNDM2	1,330.43	686.93	471.70	362.96	300.23	256.32	242.10	243.20
FSBNDM	1,273.48	667.56	460.61	357.41	294.52	252.65	242.10	241.95
GSB2	1,349.88	694.35	479.60	367.18	302.10	259.52	245.15	247.73
SBNDM4	1,195.82	621.93	427.41	328.58	272.26	232.81	221.17	221.79
SBNDM6	1,347.38	543.34	379.91	304.60	255.07	220.46	208.51	208.67
SBNDM8	1,353.48	547.09	382.72	304.68	258.12	220.85	209.60	208.82
SSM	1841.74	1589.48	1502.06	1408.16	1370.90	1344.73	1286.37	1297.85
FSB (3,1)	1,219.34	633.97	438.82	338.35	281.09	241.63	231.71	232.18
FSB (3,2)	1,246.29	636.00	440.61	339.91	281.01	241.09	235.31	235.62
FSB (4,1)	1,208.17	619.98	428.66	331.08	274.52	234.13	225.62	225.38
FSB (4,2)	1,209.26	608.97	421.08	328.12	270.93	233.20	226.17	226.87
FSB (6,1)	1,271.06	560.76	396.00	313.35	262.41	224.84	215.70	215.77
FSB (6,2)	1,188.71	561.31	396.94	315.77	261.16	223.98	217.18	217.65
FSB (6,4)	1,283.48	610.29	409.75	313.51	256.16	217.73	218.27	217.88
FSB (8,1)	1,997.97	596.92	364.13	268.33	218.56	188.15	180.40	180.79
FSB (8,2)	1,637.92	566.56	355.54	264.53	217.18	187.26	182.11	182.34
FSB (8,3)	1,427.31	549.91	351.94	264.60	218.12	188.04	185.78	185.70
HASH2	1,681.36	1,683.31	1,666.04	1,651.20	1,695.65	1,672.61	1,653.54	1,645.89
HASH3	1,133.25	854.66	814.19	816.38	795.45	777.71	771.93	806.07
HASH4	1,008.72	561.55	468.03	427.02	411.71	410.61	405.22	401.47
HASH6	1,381.68	495.14	323.66	251.09	214.45	193.28	178.51	168.75
HSSM2	1,583.94	1,145.67	962.78	892.55	825.83	789.82	757.17	744.51
HSSM3	1,116.14	724.82	607.72	558.81	504.83	486.94	464.05	443.74
HSSM4	954.89	530.22	404.91	364.91	327.02	316.24	295.30	283.98
HSSM6	1,090.36	407.64	259.29	201.71	171.95	154.06	143.82	136.40

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Performance of HSSMq on random binary dataset and short patterns

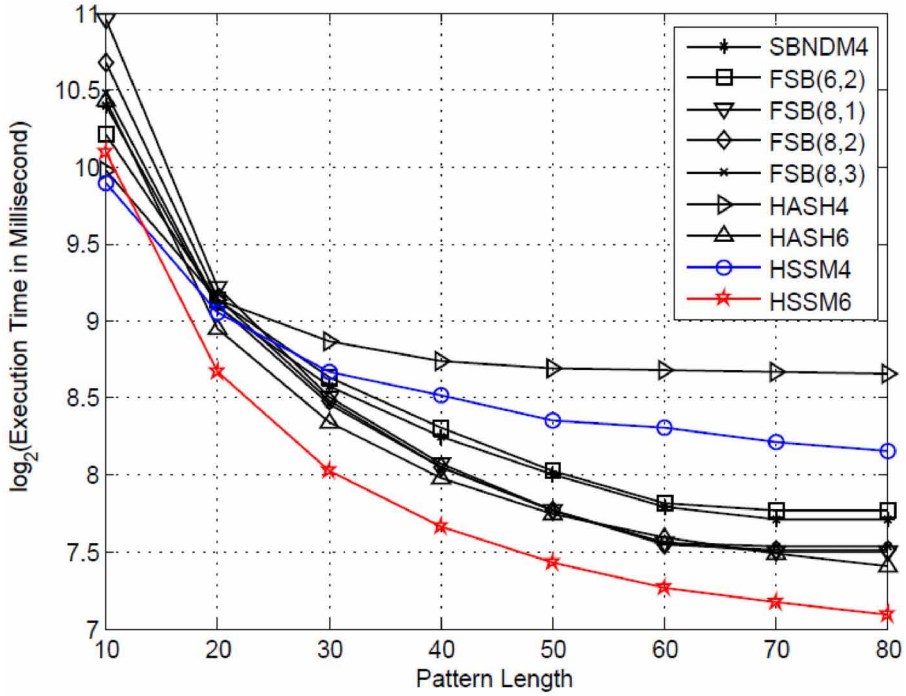
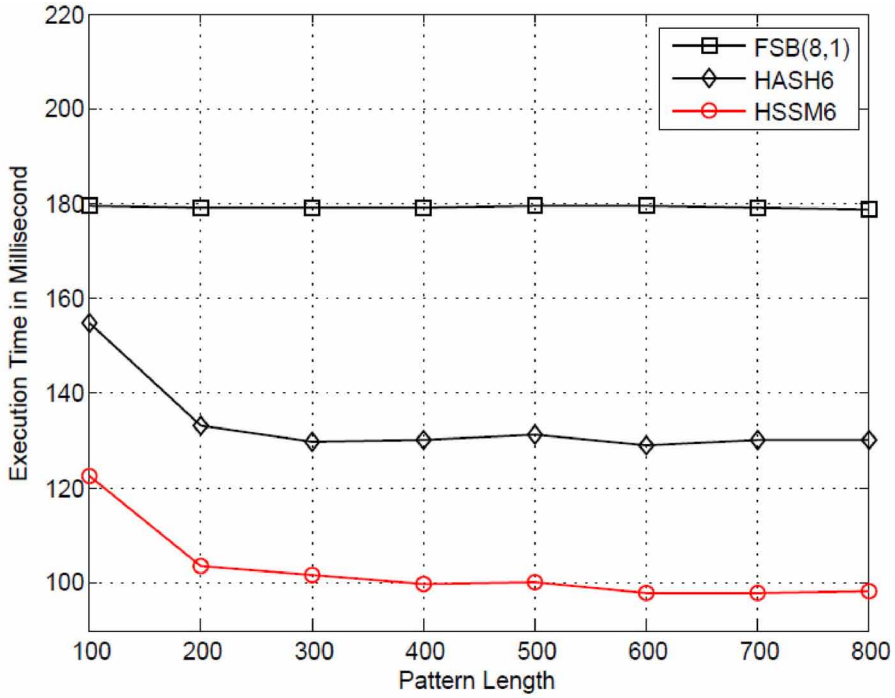


Figure 10. Running times on random binary dataset and long patterns



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. The average running times on random binary dataset and long pattern lengths

	100	200	300	400	500	600	700	800
SBNDM2	241.24	241.40	253.43	250.23	246.56	247.10	244.21	241.09
FSBNDM	240.15	240.62	239.99	240.54	240.07	239.99	239.76	240.38
GSB2	245.70	245.07	245.93	245.77	246.09	245.23	245.15	245.07
SBNDM4	222.34	220.93	220.77	222.02	220.93	220.93	221.24	221.17
SBNDM6	206.09	206.48	206.24	206.09	206.09	206.09	205.70	206.40
SBNDM8	207.18	206.56	206.79	206.24	206.24	206.63	206.95	206.95
SSM	1235.90	1155.28	1089.03	1029.58	987.94	983.80	972.71	944.66
FSB (3,1)	230.93	230.62	231.56	231.32	231.48	231.71	230.93	231.95
FSB (3,2)	233.59	234.84	234.52	234.91	235.15	234.68	234.99	235.54
FSB (4,1)	225.18	225.41	224.86	225.72	224.86	225.41	225.47	224.53
FSB (4,2)	226.25	226.10	226.64	225.93	226.25	225.70	225.54	225.39
FSB (6,1)	215.00	215.23	214.92	215.85	215.15	214.76	215.70	215.31
FSB (6,2)	216.63	216.48	217.18	216.71	217.10	216.71	216.63	216.79
FSB (6,4)	218.98	219.13	218.82	218.74	219.06	219.13	218.59	219.29
FSB (8,1)	179.45	179.06	179.06	179.14	179.53	179.45	179.14	178.90
FSB (8,2)	181.17	181.48	181.32	181.40	181.40	181.40	181.56	181.40
FSB (8,3)	184.84	185.07	184.68	184.60	185.46	184.68	184.68	184.92
HASH2	1669.95	1622.92	1631.67	1625.97	1649.48	1649.09	1720.65	1647.92
HASH3	807.48	760.91	809.04	821.15	813.10	801.30	784.12	794.43
HASH4	401.71	393.50	409.36	401.63	395.07	398.97	398.50	397.64
HASH6	154.76	132.96	129.68	130.07	131.09	129.06	130.07	130.00
HSSM2	699.90	613.65	587.8	562.01	539.36	517.64	510.61	492.88
HSSM3	418.35	368.66	346.08	335.54	325.85	320.54	314.37	312.26
HSSM4	281.01	243.74	230.15	228.04	219.92	225.85	222.65	222.26
HSSM6	122.50	103.67	101.56	99.53	100.23	97.97	97.81	98.28

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Running times on random binary dataset and very long patterns

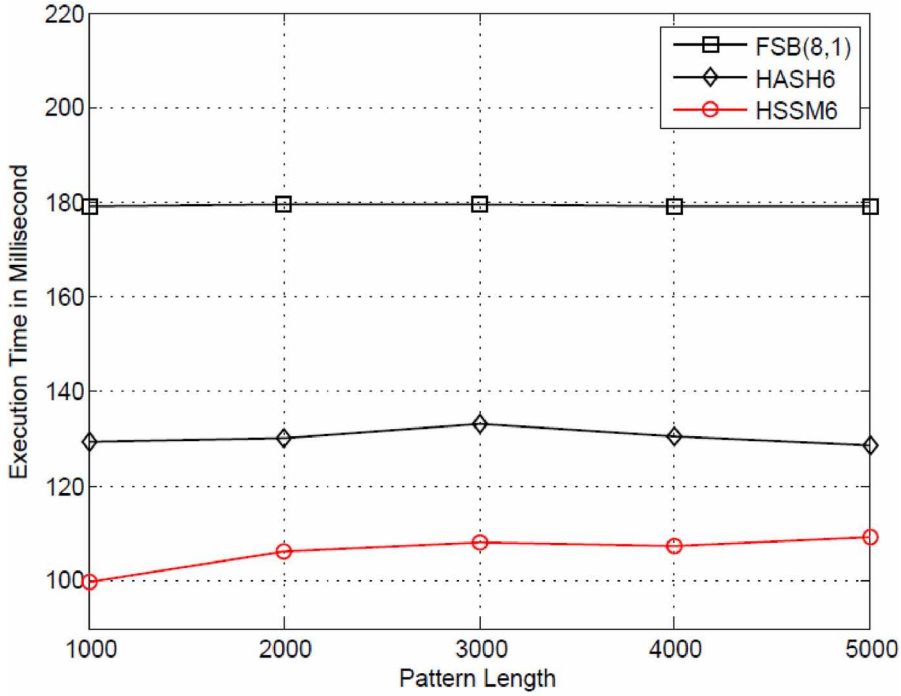
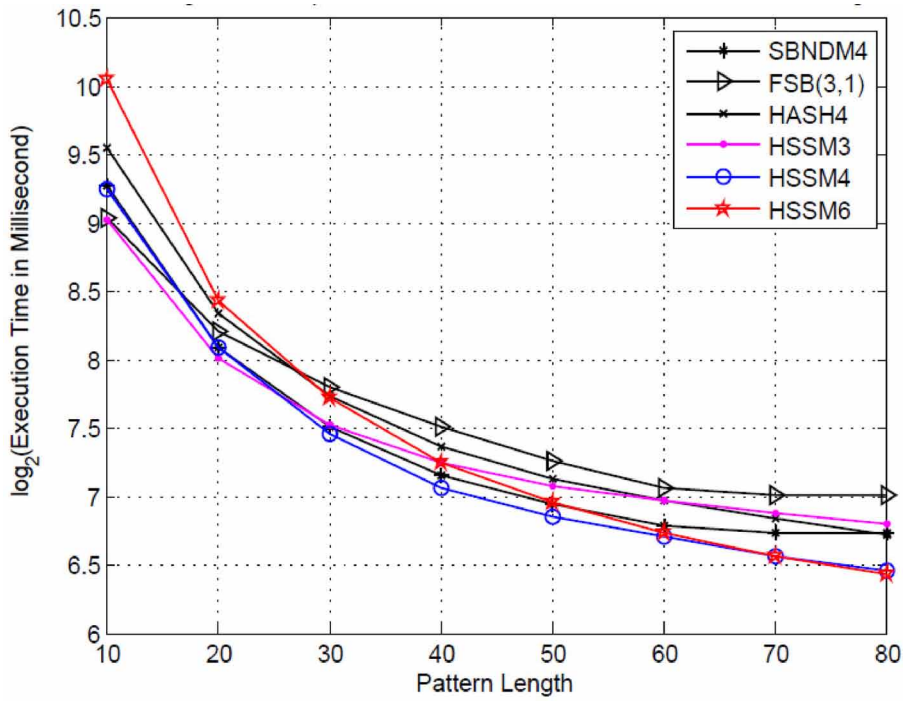


Figure 12. Running times on random DNA dataset and short patterns



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. The average running times on random binary dataset and very long patterns

	1,000	2,000	3,000	4,000	5,000
SBNDM2	240.94	242.11	241.09	241.33	242.03
FSBNDM	241.87	240.86	241.88	241.02	241.95
GSB2	244.77	245.08	244.38	245.00	244.77
SBNDM4	220.86	221.88	221.09	221.33	220.63
SBNDM6	205.78	206.09	206.80	209.45	205.86
SBNDM8	206.64	206.80	208.28	206.88	206.95
SSM	927.42	867.34	853.36	819.06	793.52
FSB(3,1)	231.80	231.33	232.03	231.41	231.17
FSB(3,2)	234.69	234.92	235.23	234.69	234.61
FSB(4,1)	225.47	225.39	225.23	225.86	226.80
FSB(4,2)	227.03	225.94	227.19	226.33	226.02
FSB(6,1)	214.84	214.69	215.30	215.30	214.83
FSB(6,2)	216.71	216.39	216.86	216.63	216.64
FSB(6,4)	218.75	218.28	218.59	218.51	218.59
FSB(8,1)	179.22	179.53	179.45	179.30	179.22
FSB(8,2)	180.94	181.09	180.78	181.09	181.41
FSB(8,3)	185.70	184.45	184.69	185.08	184.45
HASH2	1688.20	1614.45	1728.67	1757.89	1672.81
HASH3	811.80	792.42	795.63	775.08	801.56
HASH4	395.39	392.81	406.17	390.08	398.28
HASH6	129.45	130.23	133.13	130.44	128.65
HSSM2	497.95	439.29	425.60	420.52	410.60
HSSM3	304.91	283.59	282.89	276.32	273.90
HSSM4	217.73	220.93	223.04	219.14	213.44
HSSM6	99.53	106.33	107.89	107.26	109.06

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 13. Running times on random DNA dataset and long patterns

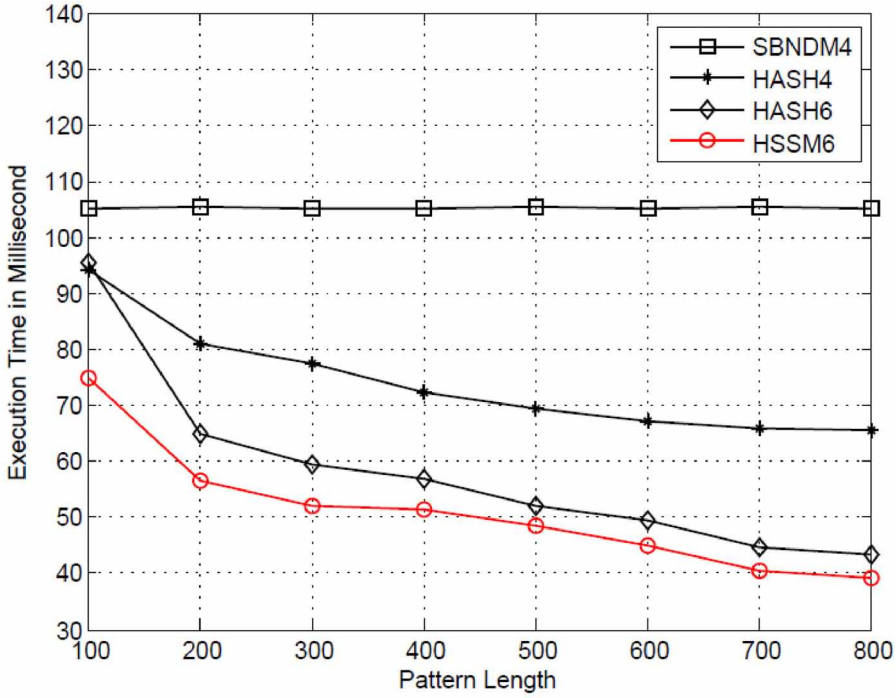
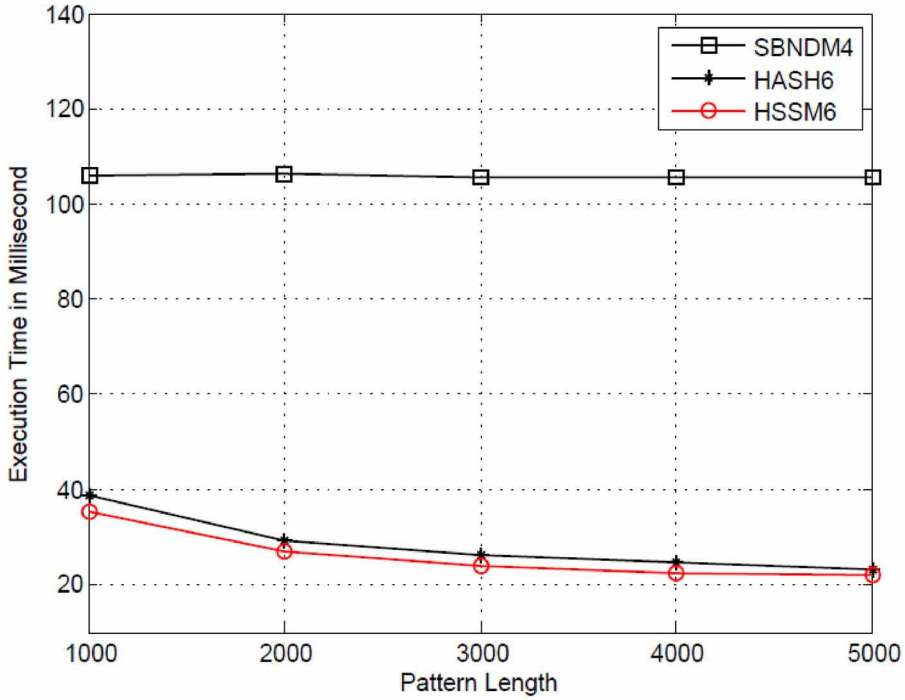


Figure 14. Running times on random DNA dataset and very long patterns



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The average running times on random DNA dataset of size 200 MB and short lengths

	10	20	30	40	50	60	70	80
SBNDM2	619.67	339.60	241.63	193.90	162.50	140.62	133.75	133.67
FSBNDM	619.67	331.63	239.91	192.03	160.93	138.82	133.59	133.90
GSB2	630.37	357.10	254.05	201.48	168.67	145.86	138.82	138.98
SBNDM4	615.87	270.61	181.35	142.65	122.33	109.67	105.90	106.03
SBNDM6	1186.06	396.74	239.93	173.52	140.87	120.47	114.54	114.54
SBNDM8	1185.70	396.95	240.07	173.43	140.78	120.31	114.45	114.45
SSM	857.79	670.76	607.01	576.70	561.31	553.50	525.22	500.22
FSB (3,1)	523.73	295.46	222.65	182.34	153.35	133.04	128.12	128.20
FSB (3,2)	637.95	345.22	240.38	185.93	154.06	133.82	131.17	130.93
FSB (4,1)	589.44	277.26	188.59	148.98	127.65	114.22	111.17	111.25
FSB (4,2)	575.14	286.09	197.10	156.25	133.59	119.14	116.87	117.03
FSB (6,1)	1066.06	402.49	249.68	182.34	148.28	127.11	122.18	122.26
FSB (6,2)	918.25	380.46	241.48	178.67	146.56	125.78	122.57	122.57
FSB (6,4)	727.09	344.76	227.10	172.81	142.26	122.97	122.97	122.97
FSB (8,1)	1933.61	554.20	324.21	230.07	181.56	152.26	145.70	145.78
FSB (8,2)	1550.11	516.94	311.08	223.35	179.29	149.92	145.86	147.03
FSB (8,3)	1301.84	492.33	304.52	219.76	176.71	148.90	146.95	146.71
HASH2	606.47	397.72	347.02	316.79	315.62	310.23	305.07	301.63
HASH3	607.09	300.23	216.40	179.29	160.78	147.34	138.67	131.79
HASH4	745.92	321.94	211.56	164.84	139.61	125.00	113.90	105.47
HASH6	1287.85	435.46	264.37	191.40	154.84	132.34	118.43	108.28
HSSM2	566.86	358.74	299.99	272.96	254.21	242.88	234.21	234.29
HSSM3	518.73	256.01	184.06	151.09	135.07	124.76	117.11	110.70
HSSM4	607.01	272.26	174.99	133.90	115.39	104.53	94.29	87.81
HSSM6	1061.53	344.44	210.62	151.71	124.14	105.78	94.29	86.33

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. The average running times on random DNA dataset and long pattern lengths

	100	200	300	400	500	600	700	800
SBNDM2	133.67	133.59	133.67	133.43	133.51	133.59	133.59	133.67
FSBNDM	133.98	134.06	133.90	134.14	133.98	133.90	134.14	134.21
GSB2	138.12	138.20	138.28	138.28	138.20	138.20	138.12	138.20
SBNDM4	105.23	105.31	105.23	105.23	105.31	105.15	105.31	105.07
SBNDM6	114.14	114.14	114.06	114.06	114.06	114.45	114.14	114.14
SBNDM8	114.29	114.22	114.22	114.22	114.14	114.29	114.22	114.14
SSM	508.19	461.78	435.22	426.24	415.61	411.32	401.47	403.97
FSB (3,1)	128.67	128.59	128.59	128.75	128.67	128.75	128.59	128.75
FSB (3,2)	131.09	131.17	131.01	130.93	131.17	131.09	131.09	131.09
FSB (4,1)	111.48	111.56	111.32	111.40	111.64	111.48	111.40	111.40
FSB (4,2)	117.42	117.18	117.18	117.26	117.26	117.26	117.42	117.34
FSB (6,1)	123.75	123.90	123.67	123.51	123.67	123.67	123.51	123.59
FSB (6,2)	123.82	123.82	123.90	123.90	123.82	123.90	123.90	123.82
FSB (6,4)	124.06	124.14	123.90	123.98	123.98	123.98	124.14	123.98
FSB (8,1)	148.75	148.83	148.76	148.75	148.76	148.91	148.76	148.83
FSB (8,2)	148.99	148.93	149.09	149.01	149.01	149.01	149.01	149.09
FSB (8,3)	149.09	148.99	148.99	148.91	148.91	148.99	149.07	149.15
HASH2	296.10	298.44	299.30	299.76	300.00	296.56	298.43	300.78
HASH3	123.98	112.11	110.86	110.07	108.98	108.90	109.06	108.67
HASH4	94.06	81.01	77.34	72.34	69.37	67.26	66.01	65.47
HASH6	95.39	64.76	59.53	56.87	52.03	49.37	44.69	43.28
HSSM2	222.88	205.38	197.26	195.78	189.84	194.45	190.62	197.18
HSSM3	104.06	92.58	89.84	89.14	88.28	87.73	86.56	87.42
HSSM4	77.65	70.94	68.90	64.37	61.56	59.76	58.90	58.12
HSSM6	74.92	56.48	52.03	51.40	48.44	45.08	40.39	39.30

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. The average running times on random DNA dataset and very long patterns

	1,000	2,000	3,000	4,000	5,000
SBNDM2	133.67	133.67	133.75	133.67	134.22
FSBNDM	136.17	136.17	133.67	133.91	133.83
GSB2	137.81	137.89	137.89	138.13	137.81
SBNDM4	106.02	106.17	105.47	105.55	105.55
SBNDM6	114.77	114.61	114.61	114.77	114.77
SBNDM8	114.53	114.84	115.39	114.53	114.84
SSM	386.48	363.52	354.77	344.61	347.03
FSB (3,1)	128.36	128.36	128.13	128.28	128.20
FSB (3,2)	131.56	131.72	131.56	131.64	131.56
FSB (4,1)	111.48	111.56	111.56	111.72	111.48
FSB (4,2)	117.03	116.95	116.95	117.19	116.95
FSB (6,1)	122.73	122.19	122.19	122.19	122.11
FSB (6,2)	122.58	122.58	122.42	122.58	122.42
FSB (6,4)	122.66	122.42	122.66	122.58	122.58
FSB (8,1)	145.78	145.78	145.63	145.63	145.55
FSB (8,2)	145.94	145.70	145.78	145.78	145.78
FSB (8,3)	145.86	145.78	145.78	145.94	145.78
HASH2	297.58	293.28	302.58	309.14	299.22
HASH3	109.14	108.98	108.91	109.45	109.61
HASH4	65.23	64.92	64.53	64.53	65.47
HASH6	38.67	29.30	26.17	24.61	23.28
HSSM2	209.44	218.19	211.00	207.56	205.92
HSSM3	88.74	95.15	98.82	98.82	98.20
HSSM4	57.26	56.25	56.01	56.25	56.40
HSSM6	35.39	26.87	23.75	22.34	21.95

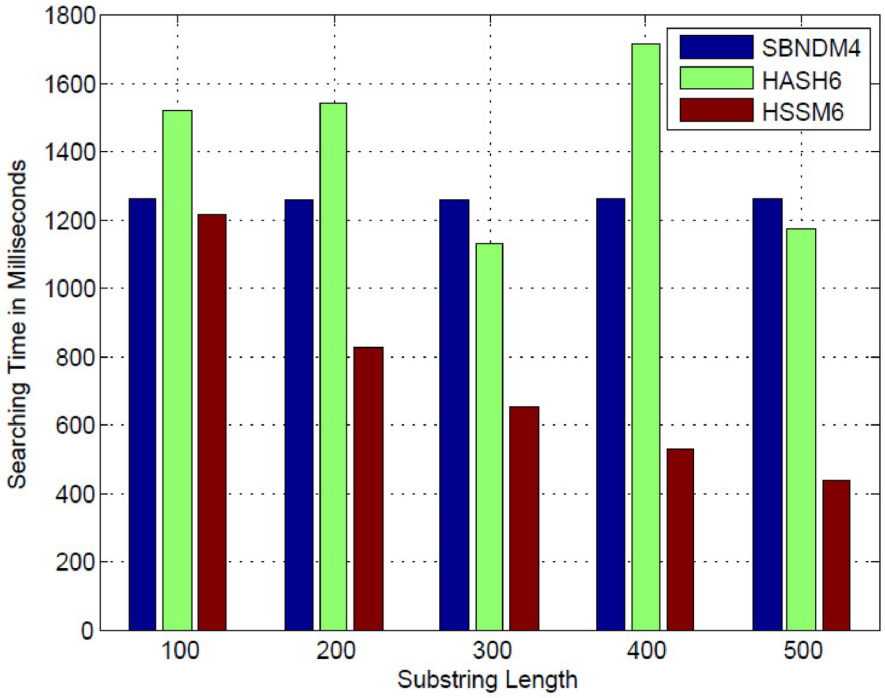
We test the algorithms, SBNDM4, HASH6, and HSSM6 on complete Homo sapiens genome (Assembly GRCh38.p3, Accession GCA_000001405.18)[2]. The elapsed time is computed for the preprocessing, searching, and reading phases and averaged over 500 randomly generated patterns. The results in Figure 15 show that HSSM6 is the most efficient algorithm. HASH6 exhibited unstable running time because the long runs of A's and T's in the genome where the algorithm shifts slower. HSSM6 overcomes this situation.

Overall, we can say that the experimental results prove that our new algorithm HSSMq performs the searching process faster than the previous algorithms.

The significance of this research comes from the importance of the two applications Intrusion Detection Systems and bioinformatics. In Intrusion Detection Systems, this algorithm can be adapted for inspecting the network streams for detecting predefined malwares while in bioinformatics it will be used for searching bioinformatics databases for a particular DNA segment. Since string matching is used frequently on these applications, the impact will be great on the performance of the two applications.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. Searching times on complete human genome



5. CONCLUSION AND FUTURE WORK

An efficient algorithm for binary and DNA texts is presented. The new algorithm combines the merits of HASHq and SSM algorithms. The new algorithm uses the q-gram hashing and guarantees the maximal shift in advance by determining the best q-gram pivot. The searching phase of the algorithm matches the hash value of the q-gram pivot with the hash value of the corresponding q-gram text window. If they are equal, a verification step is performed and the new shift is applied; otherwise, the HASHq shift is applied.

We conducted the experiments on two types of random data, binary and DNA. Then we measured the average shift length of HSSMq against HASHq and we found that there is a clear improvement on the length of shift on both random data, DNA and binary. In addition, we run the algorithms on different lengths and the results demonstrate that the new algorithm is faster than SSM, HASHq, SBNDMq and $FSB(q,f)$ for short, long and very long patterns. Regarding the binary data, HASH6 achieved better performance than the previous algorithms and the new algorithm HSSM6 outperformed HASH6 so that the amount of saved running time was about 18-21% for pattern lengths 10-80, 21-25% for lengths 100-800, and reduced by about 18-23% on lengths 1,000-5,000. On the DNA text, HSSM6 achieved a better performance than all the others. For long patterns (100-800), HSSM6 reduced the running time of HASH6 by about 7-21% and for very long patterns (1,000-5,000), the amount of saved time was 6-9%.

At the end, we run the two algorithms on the complete human genome. The results show that our algorithm is faster. HSSMq searches the whole genome for pattern of length 500 in 432 milliseconds, HASH6 spends 1170 milliseconds, and SBNDM4 takes 1258 milliseconds. So, we recommend to adapting the new algorithm for the Intrusion Detection Systems and bioinformatics.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As future research study, this work can be extended for the following topics: two-dimensional pattern matching for image processing applications, multiple pattern matching in which a matching process is performed on a set of patterns simultaneously, approximate string matching, pattern matching with wildcards, and circular string matching problems.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests regarding the publication of this paper.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through Research Group no. RGP-1436-002.

REFERENCES

Al-Ssulami, A. M. (2015). Hybrid string matching algorithm with a pivot. *Journal of Information Science*, *41*(1), 82–88. doi:10.1177/0165551514555668

Allauzen, C., Crochemore, M., & Raffinot, M. (1999). Factor Oracle: A New Structure for Pattern Matching. *Paper presented at the 26th Conference on Current Trends in Theory and Practice of Informatics on Theory and Practice of Informatics*. 10.1007/3-540-47849-3_18

Allauzen, C., & Raffinot, M. (2000). Simple Optimal String Matching Algorithm. *Journal of Algorithms*, *36*(1), 102–116. doi:10.1006/jagm.2000.1087

Baeza-Yates, R., & Gonnet, G. H. (1992). A new approach to text searching. *Communications of the ACM*, *35*(10), 74–82. doi:10.1145/135239.135243

Baeza-Yates, R. A. (1989). Improved string searching. *Software, Practice & Experience*, *19*(3), 257–271. doi:10.1002pe.4380190305

Barton, C., Iliopoulos, C., & Pissis, S. (2014). Fast algorithms for approximate circular string matching. *Algorithms for Molecular Biology; AMB*, *9*(1), 1–10. doi:10.1186/1748-7188-9-9 PMID:24656145

Blumer, A., Blumer, J., Haussler, D., Ehrenfeucht, A., Chen, M. T., & Seiferas, J. (1985). The smallest automation recognizing the subwords of a text. *Theoretical Computer Science*, *40*, 31–55. doi:10.1016/0304-3975(85)90157-4

Boyer, R. S., & Moore, J. S. (1977). A fast string searching algorithm. *Communications of the ACM*, *20*(10), 762–772. doi:10.1145/359842.359859

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Branislav, D. (2010). Improving practical exact string matching. *Information Processing Letters*, *110*(4), 148–152. doi:10.1016/j.ipl.2009.11.010

Cantone, D., Faro, S., & Giaquinta, E. (2010). A compact representation of nondeterministic (suffix) automata for the bit-parallel approach. *Paper presented at the Annual Symposium on Combinatorial Pattern Matching*. 10.1007/978-3-642-13509-5_26

Chen, K.-H., Huang, G.-S., & Lee, R. C.-T. (2014). Bit-parallel algorithms for exact circular string matching. *The Computer Journal*, *57*(5), 731–743. doi:10.1093/comjnl/bxt023

Chu, Y.-M., Huang, N.-F., Tsai, C.-H., & Hsieh, C.-Y. (2008). A software-based string matching algorithm for resource-restricted network system. *Communications Letters, IEEE*, *12*(8), 599–601. doi:10.1109/LCOMM.2008.080253

Deusdado, S., & Carvalho, P. (2009). GRASPm: An efficient algorithm for exact pattern matching in genomic sequences. *International Journal of Bioinformatics Research and Applications*, *5*(4), 385–401. doi:10.1504/IJBRA.2009.027510 PMID:19640827

Fan, H., Yao, N., & Ma, H. (2009, 21-22 Dec. 2009). *Fast Variants of the Backward-Oracle-Marching Algorithm.* Paper presented at the Internet Computing for Science and Engineering (ICICSE), 2009 Fourth International Conference on.

Faro, S., & Lecroq, T. (2009a). An Efficient Matching Algorithm for Encoded DNA Sequences and Binary Strings. *Paper presented at the 20th Annual Symposium on Combinatorial Pattern Matching*, Lille, France. 10.1007/978-3-642-02441-2_10

Faro, S., & Lecroq, T. (2009b). Efficient variants of the Backward-Oracle-Matching algorithm. *International Journal of Foundations of Computer Science*, *20*(06), 967–984. doi:10.1142/S0129054109006991

Faro, S., & Lecroq, T. (2012, November 11-13). Fast searching in biological sequences using multiple hash functions. *Paper presented at the 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE).*

Faro, S., & Lecroq, T. (2013). The exact online string matching problem: A review of the most recent results. *ACM Computing Surveys*, *45*(2), 1–42. doi:10.1145/2431211.2431212

Franek, F., Jennings, C. G., & Smyth, W. F. (2007). A simple fast hybrid pattern-matching algorithm. *Journal of Discrete Algorithms*, *5*(4), 682–695. doi:10.1016/j.jda.2006.11.004

He, L., Fang, B., & Sui, J. (2005). The wide window string matching algorithm. *Theoretical Computer Science*, *332*(1-3), 391–404. doi:10.1016/j.tcs.2004.12.002

Horspool, R. N. (1980). Practical fast searching in strings. *Software, Practice & Experience*, *10*(6), 501–506. doi:10.1002pe.4380100608

James, H., Morris, J., & Pratt, V. R. (1970). *A linear pattern-matching algorithm Technical Report*. Berkeley: University of California.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kalsi, P., Peltola, H., & Tarhio, J. (2008). Comparison of Exact String Matching Algorithms for Biological Sequences. In M. Elloumi, J. Küng, M. Linial, R. Murphy, K. Schneider, & C. Toma (Eds.), *Bioinformatics Research and Development* (Vol. 13, pp. 417–426). Springer Berlin Heidelberg. doi:10.1007/978-3-540-70600-7_31

Knuth, D. E., James, H., Morris, J., & Pratt, V. R. (1977). Fast Pattern Matching in Strings. *SIAM Journal on Computing*, *6*(2), 323–350. doi:10.1137/0206024

Lecroq, T. (2007). Fast exact string matching algorithms. *Information Processing Letters*, *102*(6), 229–235. doi:10.1016/j.ipl.2007.01.002

Liu, R.-T., Huang, N.-F., Chen, C.-H., & Kao, C.-N. (2004). A fast string-matching algorithm for network processor-based intrusion detection system. *ACM Transactions on Embedded Computing Systems*, *3*(3), 614–633. doi:10.1145/1015047.1015055

Mohanty, P., & Tragoudas, S. (2014). Scalable Offline Searches in DNA Sequences. *J. Emerg. Technol. Comput. Syst.*, *11*(2), 1–25. doi:10.1145/2660774

Morozova, O., & Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics*, *92*(5), 255–264. doi:10.1016/j.ygeno.2008.07.001 PMID:18703132

Navarro, G., & Raffinot, M. (2000). Fast and flexible string matching by combining bit-parallelism and suffix automata. *J. Exp. Algorithmics*, *5*, 4, es. doi:10.1145/351827.384246

Peltola, H., & Tarhio, J. (2003). Alternative Algorithms for Bit-Parallel String Matching. In M. Nascimento, E. de Moura, & A. Oliveira (Eds.), *String Processing and Information Retrieval* (Vol. 2857, pp. 80–93). Springer Berlin Heidelberg. doi:10.1007/978-3-540-39984-1_7

Peltola, H., & Tarhio, J. (2014). String matching with lookahead. *Discrete Applied Mathematics*, *163*(Part 3), 352–360. doi:10.1016/j.dam.2013.10.034

Simon, I. (1994). String Matching Algorithms and Automata. *Paper presented at the Colloquium in Honor of Arto Salomaa on Results and Trends in Theoretical Computer Science*. 10.1007/3-540-58131-6_61

Sunday, D. M. (1990). A very fast substring search algorithm. *Communications of the ACM*, *33*(8), 132–142. doi:10.1145/79173.79184

Tarhio, J., & Peltola, H. (1997). String matching in the DNA alphabet. *Software, Practice & Experience*, *27*(7), 851–861. doi:10.1002/(SICI)1097-024X(199707)27:7<851::AID-SPE108>3.0.CO;2-D

Tuck, N., Sherwood, T., Calder, B., & Varghese, G. (2004). Deterministic memory-efficient string matching algorithms for intrusion detection. *Paper presented at the INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*. 10.1109/INFCOM.2004.1354682

Wu, S., & Manber, U. (1992). Fast text searching: Allowing errors. *Communications of the ACM*, *35*(10), 83–91. doi:10.1145/135239.135244

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] http://www.dmi.unict.it/~faro/smart/

[2] www.ensembl.org

This research was previously published in the International Journal on Semantic Web and Information Systems (IJSWIS), 13(4); edited by Miltiadis D. Lytras; pages 198-220; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 17
Concept Attribute Labeling and Context-Aware Named Entity Recognition in Electronic Health Records

Alexandra Pomares-Quimbaya
Pontificia Universidad Javeriana, Colombia

Angel Alberto Garcia Peña
Pontificia Universidad Javeriana, Colombia

Rafael A. Gonzalez
Pontificia Universidad Javeriana, Colombia

Julián Camilo Daza Rodríguez
Pontificia Universidad Javeriana, Colombia

Oscar Mauricio Muñoz Velandia
Pontificia Universidad Javeriana, Colombia

Alejandro Sierra Múnera
Pontificia Universidad Javeriana, Colombia

Cyril Labbé
University Grenoble Alpes, France

ABSTRACT

Extracting valuable knowledge from Electronic Health Records (EHR) represents a challenging task due to the presence of both structured and unstructured data, including codified fields, images and test results. Narrative text in particular contains a variety of notes which are diverse in language and detail, as well as being full of ad hoc terminology, including acronyms and jargon, which is especially challenging in non-English EHR, where there is a dearth of annotated corpora or trained case sets. This paper proposes an approach for NER and concept attribute labeling for EHR that takes into consideration the contextual words around the entity of interest to determine its sense. The approach proposes a composition method of three different NER methods, together with the analysis of the context (neighboring words) using an ensemble classification model. This contributes to disambiguate NER, as well as labeling the concept as confirmed, negated, speculative, pending or antecedent. Results show an improvement of the recall and a limited impact on precision for the NER process.

DOI: 10.4018/978-1-7998-1204-3.ch017

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Electronic health records (EHR) constitute an important resource not just for tracing single patient histories but for population studies with clinical or administrative purposes. The nature of EHR, however, presents multiple challenges for doing so. In fact, physicians often complain that EHR systems are oriented towards information storage, but lack the ability to provide information extraction as well, mainly a consequence of the unstructured nature of the information (Menasalvas, Rodriguez-Gonzalez, Costumero, Ambit, & Gonzalo, 2016). Actually, the extraction task involves a combination of structured and unstructured data, including codified clinical classifications, images, test results and narrative text, among others. This paper will focus on the recognition of pre-established entities of interest in EHR extracted from clinical systems. This falls within the task of named entity recognition (NER), responsible for extracting entities and relationships between them, within a specific domain, typically relying on dictionaries, ontologies and thesaurus to do so (Menasalvas et al., 2016). As such, this paper proposes an approach to NER, which is aimed at improving precision and recall by combining different NER methods.

There is a large body of work on methods and tools to process biomedical text in general (Menasalvas et al., 2016). However, as pointed out in Leaman et.al (Leaman, Khare, & Lu 2015), biomedical texts are a highly codified result edited for clarity and intended at a large audience, while clinical narrative texts contained in EHR are written by healthcare professionals about a single patient and are aimed at colleagues or themselves. This implies a variety of notes, which are diverse in language and detail, as well as ad hoc terminology, including acronyms and jargon, far from being highly codified and standard. In practice, these results in EHR systems are country, hospital and even service dependent (Menasalvas et al., 2016). In addition, EHR are often filled under time pressure and with low motivation due to the fact that it takes time away from actual patient care or education. As a result, EHR narrative text usually suffers from low quality reflected in: variable semantics, structure without formal sentences, missing punctuation, missing expected words, misspelling or heterogeneous styles and jargon (Menasalvas et al., 2016). Moreover, independently of the motivation or resulting quality, the clinical language per se implies an additional series of challenges, including term variability, ambiguity and complexity, lack of fine-grained classifications, results followed by units or dosages, incomplete syntactic components in sentences, as well as data availability (Dong, Qian, Guan, Huang, Yu, & Yang, 2016). In NER terms, ambiguity is one of the biggest challenges, because concepts of interest are frequently hypothetical, negated or include temporal relationships (Menasalvas et al., 2016). As such, many existing natural language processing (NLP) approaches become ineffective or insufficient for clinical narrative text.

Moreover, despite numerous NER proposals for EHR, the vast majority are limited to medical text written in English (Menasalvas et al., 2016). Given that NER relies on dictionaries, several are already available, including Unified Medical Language System (UMLS), Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT) or International Classification of Diseases (ICD); these are also in English with either no translations or limited versions in other languages. In the context of this paper, EHRs belong to a Spanish speaking hospital and it has already been recognized that for event extraction from EHRs in Spanish, the lack of annotated corpora is perhaps the main difficulty (Casillas, Pérez, Oronoz, Gojenola, & Santiso, 2016). Despite there being ways to avoid a language-specific annotated corpus through supervised machine learning, training the models is costly and using it in complex language sets runs into major performance issues (Dong et al., 2016). In addition, the choice of inference algorithms and managing heterogeneous medical fields further complicates medical NER (Casillas et al., 2016; Dong et al., 2016).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This paper proposes an approach for dictionary-based, combined NER aimed at improving entity recall dealing with the aforementioned challenges associated to clinical narratives extracted from EHR systems. To do so, it presents a proposal that combines three different NER techniques and takes into consideration the contextual words around the entity of interest to determine whether the concept is confirmed, negated, speculative, pending or an antecedent. To our best knowledge, our approach is the first including this level of detail in the expression of sense for the healthcare domain in Spanish. The rest of this paper continues with related works. We then go on to present the proposed strategy and present the evaluation results of applying the strategy to a standard data set. Finally, we present some conclusions and suggests avenues for future work.

BACKGROUND

Natural language processing of medical text has a wide variety of tools and methods. For our research work, we made an analysis of some relevant work in NER in the medical and biomedical fields in Table 1, where we proceed to highlight the following characteristics:

- **NER Technique:** This characteristic summarizes the strategy used for identifying the entity;
- **Entities:** Indicates the type of concepts the strategy is able to recognize;
- **Features:** Illustrates whether the approach includes only the entity recognition or other attributes around the entity (e.g. negation);
- **Limitations:** This characteristic highlights the main limitations of the approach.

Based on the previous comparison, we can find no general method for finding named medical entities, which can deal with the ambiguity and complexity of EHR discussed in the introduction. Some of the extant contributions are focused on specific entities (Jiang et al., 2011; Jimeno et al., 2008; Aramaki et al., 2009; Skeppstedt et al., 2014), some contributions (Chang et al., 2002; Zhang et al., 2013; Xu et al., 2012) lack the ability to find entities taking into account the text quality challenges described previously, others (Henriksson et al., 2015; Casillas et al., 2016) analyze the context only for a few entities in the clinical domain. Although there are recent approaches for recognizing the attributes around an entity, they are only available for languages different than Spanish and do not consider the quality problems of the texts that hinder the recognition of the entities. Our proposed approach aims to perform NER in medical texts contained in electronic health records, considering their context and their intrinsic quality or ambiguity. The following section explains in detail this approach.

CONTEXT AWARE NER FOR ELECTRONIC HEALTH RECORDS

In order to extract meaningful information from an analysis over EHR, we require not only the recognition of named entities, but also the identification of the sense they have within the text, since the entity could be negated in a specific context, for example. Moreover, disambiguating and determining the precise use of the entity in context is especially challenging for non-English medical text, where dictionaries and annotated training corpora are scarce or incomplete. In consequence, we propose a context-aware combined NER approach for EHR in Spanish (Quimbaya, Múnera, Rivera et al., 2016). This approach

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Comparative analysis of NER methods in medical text

	NER Technique	Entities	Characteristics	Limitations
Chang, Schutze, & Altman, 2002	Pre-defined Dictionary	Acronyms and initials for health information resources, Human genome acronyms	Entity Recognition Search Abbreviations	Limited to abbreviations and acronyms
Jiang, Chen, Liu et al., 2011	Conditional Random Fields CRF	Medical Problems, Test, Treatments, Status	Entity Recognition	Needs a model training process
Jimeno, Jimenez-Ruiz, Lee et al., 2008	Exact Matching Flexile Matching	Pathologic Functions, Sign or Symptom, Cell or Molecular Dysfunctions, Findings	Entity Recognition	NER is limited to scientific texts
Aramaki, Miura, Tonoike et al., 2009	Conditional Random Fields CRF	Remedy, Medical Operations, Test, Examinations, Diseases, Symptoms, Medications	Entity Recognition Events Recognitions Date Times Negative Events	Needs a model training process
Zhang & Elhadad, 2013	Seed Term Collection Boundary Detection Entity Classification	Disorders, Therapeutic or Preventive Procedures, Laboratory Procedures, Clinical Drug, Test	Entity Recognition	Can't recognize typing and orthographic errors
Skeppstedt, Kvist, Nilsson et al., 2014	Conditional Random Fields CRF Lemmatization Part-of-speech tagging Terminology match	Disorders, Findings, Pharmaceutical, Body Structure	Entity Recognition NER in English and Swedish medical texts Errors recognition	Needs a model training process Previous annotation process developed by Physician or Computational linguist
Xu, Tsujii, Eric et al., 2012	CRF-based NER Labeled sequential pattern LSP	Exams, Impressions, Conclusions	Entity Recognition Recognition of sentences with follow-up information	Needs a model training process
Henriksson, Dalianis & Kowalski, 2014	Conditional Random Fields CRF Tokenization Part-of-speech tagging	Protected Health Information PHI (First Name, Last Name, Age, Health care unit, Location, Phone Number)	Entity Recognition	Limited to identify protected health information PHI in EHR
Henriksson, Kvist, Dalianis et al., 2015	Tokenization Lemmatization and Part to Speech Terminology matching (SNOMED CT)	Drugs, Disorders, Finding, Body Structure	Entity recognition Attributes of entities (negation, speculation, future, Other) Relations between Entities (Indications, Adverse Drug Event) Unsupervised method	NER is limited to drug findings
Casillas, Pérez, Oronoz et al., 2016	Lemmatization Random Forest (RF) Support Vector Machine	Drugs, Brand Names, Active Ingredients Disease, Disorders, Symptoms, Allergies	Entity Recognition Event Extraction	Needs a model training process

first identifies the occurrence of important clinical entities such as procedures, medications, diagnosis, adverse events, among others, using a combined dictionary approach, and then analyses the context of the identified entity for detecting its sense using a machine learning approach. Our proposal allows the identification of six different senses associated to entities within EHR text, accordingly: confirmed, negated, speculative, pending, patient antecedent and family antecedent. To the best of our knowledge, our approach is the first to include this level of detail in the expression of sense for the healthcare domain in Spanish.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1 illustrates the phases included in our approach. The first phase recognizes the entities from the text using a dictionary strategy that combines three types of NER techniques and generates a preliminary tag over the relevant entities. The second phase, extracts the context of each one of the entities and generates a set of n-grams that are used as the input for the final phase that applies an ensemble classification model that enriches the preliminary tag including the sense of the entity within the EHR. The following sections explain each phase in detail.

Phase 1: Context Aware NER for Electronic Health Records

Figure 2a and 2b present the strategy proposed to improve the recall of NER in EHR implemented in the first phase of our approach. To deal with possible quality issues in the texts it follows two steps: the first one executes preprocessing tasks (Figure 2a), and the second one performs the NER including three different types of recognition (Figure 2b): exact NER, fuzzy NER and stemmed NER.

The preprocessing phase extracts the lemmas from the domain knowledge base. This task generates a new knowledge base with the lemmas of the concepts that have to be recognized. In addition, it performs preprocessing tasks including tokenization and lemmatization over the EHR.

The objective of the NER phase is to recognize the relevant entities (e.g. diagnosis, treatment, procedures, medical labs), even if they have misspellings or are written with differences with respect to the concepts included in the domain knowledge base. As mentioned, it then uses three different NER techniques. Finally, it includes a merging process that combines the overlapped recognized entities.

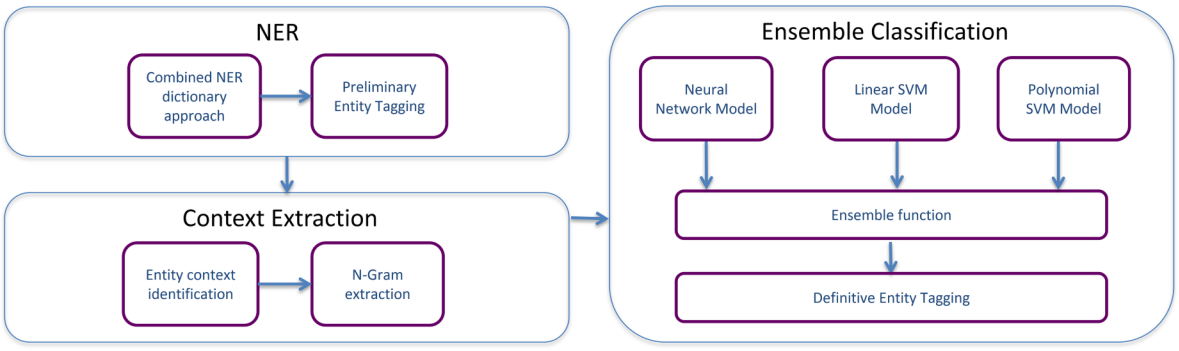
One to One Recognition

The most basic approach to NER when using a dictionary or gazetteer is simply finding exact matches between dictionary and EHR entities. That means having a one to one match between a dictionary term and a portion of the text.

For instance let G be a diagnosis dictionary, and $DocX$ a document in the corpus:

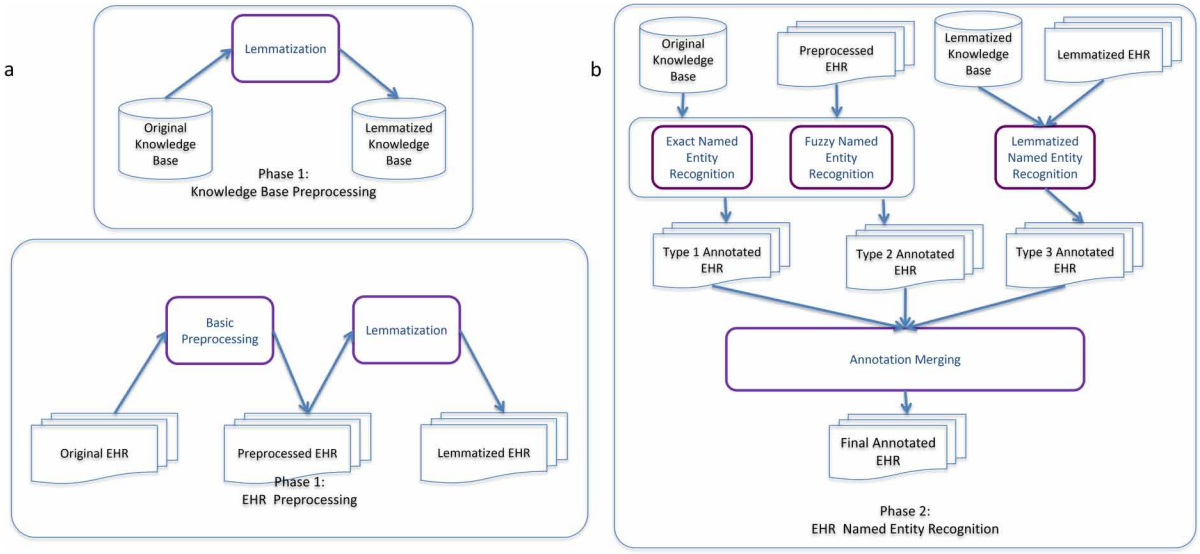
$$G = \{'diabetes','hypertension','COPD','influenza'\} \tag{1}$$

Figure 1. Context aware NER for electronic health records



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Phase 1 combined dictionary-based NER: (a) Preprocessing phase; (b) NER recognition phase



DocX = 'John Doe is a 67 year-old diabetic white male with a history of COPD, and hypertension. Mr. Doe was hospitalized 20 days ago at High Plains Hospital for pneumonia resulting from influezna ...'

Running NER with the dictionary would result in an annotated document DocXAnnot:

DocXAnnot = 'John Doe is a 67 year-old diabetic white male with a history of [**COPD**], and [**hypertension**]. Mr. Doe was hospitalized 20 days ago at High Plains Hospital for pneumonia resulting from influenza …'

This annotated document contains references for the terms 'COPD' and 'hypertension' and such annotations can be used to flag this document as relevant for such medical terms.

Fuzzy Recognition

EHRs are prone to containing typing and spelling errors. That degrades the recall of exact dictionary-based NER, due to the likelihood of a mismatch between the dictionary and the actual words in the documents. For instance, a dictionary term 'pneumonia' will not match a mistyped 'npeumonia'.

We propose using a Fuzzy Gazetteer approach (Paleo, B. W. 2007) in order to find more instances of the desired named entities, when there may be slight differences between the dictionary term and the EHR term.

In order to compare EHR terms and dictionary terms we make use of the concept of edit distance (Levenshtein, 1966) which is a metric representing the number of character changes among words. We define the threshold for accepting annotations as a proportion to the length of the dictionary term. For instance, we admit more changes in a long term like 'hypertension' than in a short one like 'COPD'. The

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

threshold parameter T ∈ [0, 1] reflects such proportion and corresponds to the maximum proportional edit distance with respect to the term´s length, as expressed in the following equation:

$$T = \frac{maximum\ edit\ distance\ accepted\ for\ a\ term}{length\ of\ \ the\ term}$$

Using the document DocX and the dictionary G with the Fuzzy Gazetteer approach, configured with a value of T = 0.25 would result in the following annotated document:

DocXAnnFuzz = 'John Doe is a 67 year-old diabetic white male with a history of [**COPD:COPD(0)**], and [**hypertension:hypertension(0)**]. Mr. Doe was hospitalized 20 days ago at High Plains Hospital for pneumonia resulting from [**influezna:influenza(2)**] ...'

Each annotation in the document has the form [text: dictionaryTerm(editDistance)] where 'text' represents the matched string in the document, 'dictionaryTerm' the term in the gazetteer and 'editDistance' the distance among them. In this example, T = 0.25, therefore the edit distance between influenza and influenza (2) divided by the term's length (9) satisfies the defined threshold:

$$\frac{2}{9} = 0.2\hat{2} < T$$

Stemmed Recognition

Recognizing entities based on the stem of a dictionary term involves having the dictionary lemmas or stems, then finding the stem of each word in the documents and then comparing the stemmed versions to exclude differences due to, for instance, plurals or conjugation. In the example, we first compute *G':* the stemmed version of the dictionary *G* and then, document *DocX* is processed with the same stemmer and as a result we have a parallel document *DocX'*:

$$G' = \{'diabet','hypertens','copd','influenza'\} \quad (2)$$

DocX'= 'john doe is a 67 year-old diabet white male with a histori of copd and hypertens mr doe was hospit 20 day ago at high plain hospit for pneumonia result from influezna ...'

Note that in *DocX'* words like 'diabetic' and 'hypertension' are respectively replaced with 'diabet' and 'hypertens'.

Once the documents and the dictionary are stemmed, they are matched and annotated. Here the annotation of 'diabet' corresponding to the stem 'diabet' shared by the words 'diabetes' and 'diabetic':

DocXAnnStem = 'john doe is a yearold [**diabet**] white male with a histori of [**copd**] and [**hypertens**] mr doe was hospit day ago at high plain hospit for pneumonia result from influezna...'

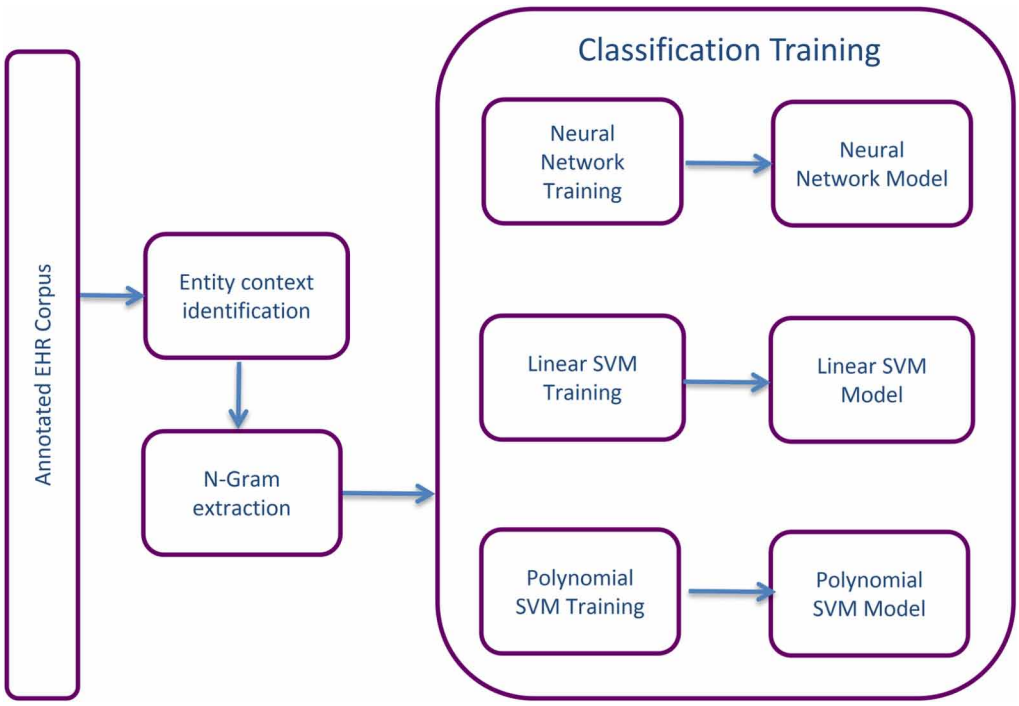
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Phase 2 and 3: Context Analysis and Entity Classification

As Figure 1 illustrates, the second phase of our proposal reads the tags included in the first phase and extracts their context, which is composed of the neighboring words that may affect the sense of the entity. Tests over a large corpus of EHR allowed us to detect that a nine-word context, seven before the entity and three after it, produced the best results in terms of precision. Nevertheless, for comparison purposes we also applied our approach using a context of four words (evaluation illustrates the results). Using the context, this phase generates a set of nine 3-gram. Subsequently, the third phase of our approach takes each one of the 3-gram to feed the ensemble model that includes three classifiers: a neural network, a linear support vector machine and a polynomial support vector machine. The combination function of the model uses a weighted voting method based on the level of confidence of the result generated by each model. Given the set of results for each model this function aggregates the confidence for each resulting class and chooses the class with the biggest aggregated confidence. For instance, the neural network classifier determines a "Negated" class with 1.0 confidence, but both the linear and polynomial SVMs agree on a "Confirmed" class with 0.9 and 0.95 confidence levels, then the final class would be "Confirmed" because the aggregated confidence for this class would be 1.85 compared to a 1.0 aggregated confidence for "Negated". That would not only work for three models as presented in our approach but could admit any number of them, so long as they include a confidence value.

The classification models used in this phase were previously trained using manually expert annotated portions of texts extracted from EHR of a general hospital. As can be seen in Figure 3, the training process identifies the contexts and generates the n-grams from the annotated texts in the same way as it is made in the general process. The annotations include six classes:

Figure 3. Sense classification training



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. **Confirmed:** This includes notes in an EHR implying a positive, affirmed presence of a specific medical condition, as well as diagnostic or treatment procedures. As an example, in the EHR it could appear a paragraph like: "Female patient hospitalized for *acute decompensated heart failure* that today refers… ". The sentence implies the confirmed presence of the disease in bold;
2. **Negated:** An entity whose presence is negated in a given context, such as "Patient with neither *erythema* nor other local infection signs suggestive of phlebitis" denies that diagnosis;
3. **Speculative:** Sometimes in EHRs the health professionals will record differential diagnosis or different therapeutic options that cannot yet be confirmed or negated and as such are registered in a speculative fashion. For example, "Patient 40 years old with first episode of dyspnea and wheezing that could be associated with *chronic obstructive pulmonary disease (COPD) exacerbation or asthmatic crisis*" is speculative for both diagnoses;
4. **Pending:** On occasion, an EHR will contain an entity related to a procedure that is planned or pending for a patient, without if being confirmed or negated, as results are still not available. For instance, "Patient with confirmed diagnosis of acute appendicitis who will be scheduled for *appendectomy* according with surgery room availability" refers to the procedure as a future event;
5. **Patient Antecedent:** Occasionally personal antecedents are not registered in structured fields on an EHR. It is risky or misleading to assume the presence of such an entity for a given query, without taking into consideration that it refers to a past condition or event which may no longer be valid. It can be seen in notes like "Patient with two previous hospitalizations in 2009 for *urinary tract infections* who …" which recalls the past and not the present condition of the patient;
6. **Familiar Antecedent:** This case is similar to the last one, with the added risk or ambiguity stemming from the fact that it does not relate to the patient directly, but to his or her family, such as, "Mother with *early myocardial infarction*".

The complete algorithm for the 3 phases is expressed as follows:

```
FUNCTION ner(document,gazetteer)
        mentions={ }
        mentions=mentions UNION findAllExactMatches(document,gazetteer)
        mentions=mentions UNION findAllFuzzyMatches(document,gazetteer)
        mentions=mentions UNION findAllStemmMatches(document,gazetteer)
        FOR EACH mention IN mentions
                senseLinearSVM=classifyLinearSVM(context(mention))
                sensePolynomialSVM=classifyPolynomialSVM(context(mention))
                sensePaum=classifyPaum(context(mention))
                sense=weightedVoting(senseLinearSVM,sensePolynomialSVM,sensePa
um)
                addAnnotation(document,mention,sense)
END
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

EVALUATION

The evaluation of our approach comprises the analysis of precision and recall measures over the NER phase as well as the precision and recall of the concept attribute labeling process. For the first part, we used a previously annotated NLP data set as the gold standard; for the second process, we developed an online software tool that facilitates the annotation process over a set of EHR. This tool allows medical doctors to classify an entity, according to the sense it has in context from the above six types or unclassified if the context is not enough to determine the sense. The results of these annotations were used as the gold standard. In both cases, the texts come from real Spanish EHR systems and they may contain typing errors, such as the ones described in the introduction. In the following sections, we present the results of our evaluation: *NER evaluation* explains the methodology and results of the combined dictionary based approach, *Data set and configuration for classification evaluation* and *Results in the classification task* illustrate details on the data set used for evaluation purposes and results of the concept attribute labeling, respectively.

NER Evaluation

The recognition of named entities was tested using the e i2b2 NLP Data Set #7b: Heart Disease Risk Factors Challenge Data Set (Stubbs, Kotfila, Xu, & Uzuner, 2015a; Stubbs, & Uzuner, 2015b). The annotations found by our approach were compared to direct mentions in the data set. The dictionary of terms was extracted from the UMLS metathesaurus (Lindberg, Humphreys, & McCray, 1993), specifically including names of diseases and medications as well as their synonyms, extracted from the MEDLINE data source. Different configurations were tested to compare the impact a direct match, fuzzy, and stemmed methods, showing an increase in recall when the fuzzy and stemmed NER annotations were introduced, while as expected the precision decreased. The test resulted in a gain of 9% in the amount of annotations, producing new information that the direct match did not include, as can be seen in Table 2. The impact of the fuzziness threshold T was also measured with a similar result in Table 3: the bigger the threshold, the bigger the amount of annotations, but after 0.15 precision becomes strongly degraded.

Data Set and Configuration for Classification Evaluation

The dataset used for training and testing the context analysis and concept attribute labeling consisted of 1477 sentences manually taken from real EHRs and manually annotated into seven classes. The six classes mentioned in 3.2: confirmed, negated, speculative, pending, patient antecedent, familiar antecedent, and an additional class N/A used for unclear fragments, which either did not have enough context or did not fit into the aforementioned classes.

Table 2. Comparison of the different approaches

	tp	fp	# Ann	fn	P	R	F1	F2
e	9124	**4750**	13874	7442	**0,658**	0,551	0,599	0,569
ef	9267	5007	14274	7299	0,649	0,559	**0,601**	0,575
es	9371	5361	14732	7195	0,636	0,566	0,599	0,578
efs	**9491**	5585	**15076**	**7075**	0,630	**0,573**	0,600	**0,583**

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Impact of the fuzziness threshold T for the ef approach

T	tp	fp	# Ann	fn	P	R	F1	F2
0,05	9146	**4768**	13914	7420	**0,657**	0,552	0,600	0,570
0,10	9210	4849	14059	7356	0,655	0,556	**0,601**	0,573
0,15	9267	5007	14274	7299	0,649	0,559	**0,601**	**0,575**
0,20	9301	5908	15209	7265	0,612	0,561	0,585	0,571
0,25	**9356**	6379	15735	7210	0,595	**0,565**	0,579	0,570
0,30	**9356**	8088	**17444**	**7210**	0,536	**0,565**	0,550	0,559

The dataset was then split into two groups, one for training the models consisting of 739 annotations and one for testing them with the remaining 738.

We also wanted to compare the nine words context with a smaller one consisting of two words before and two words after the entity. Since we are using 3-grams, the smaller context always contains the entity as part of the 3-grams group and we wanted to measure the impact of including it as input for the classification models.

The model was implemented as a GATE application both for training and testing. GATE is a well-known computer architecture for natural language processing tasks that allows to include machine learning algorithms. The testing application also had the responsibility of assembling the results of the different models into one final entity classification.

Results in the Classification Task

The results for both the long context and the short context approaches are summarized in Tables 4 and 5 as confusion matrices. Each matrix represents the amount of manual annotation for each class as the gold standard (rows) vs. the amount of annotations classified into such classes (columns). For instance, in Table 2, 190 "Negated" sentences were correctly classified as "Negated" but 1 originally "Negated" sentence was incorrectly classified as "Patient. Ant." (Second row, fifth column). The diagonal of the

Table 4. Confusion Matrix for long context classification (7 words before, 2 words after)

Gold \ Results	Confirmed	Negated	Speculative	Pending	Patient Ant.	Familiar Ant.	na	Recall
Confirmed	377	15	0	2	7	0	7	0,924
Negated	10	190	0	0	1	0	0	0,945
Speculative	15	0	0	0	0	0	0	0,000
Pending	12	0	0	1	0	0	0	0,077
Patient Ant.	40	3	0	0	3	0	3	0,061
Familiar Ant.	11	0	0	0	0	1	0	0,083
na	32	3	0	0	1	0	2	0,053
Precision	0,759	0,900	0,000	0,333	0,250	1,000	0,167	

Total Accuracy: 0.78

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Confusion Matrix for short context classification (2 words before, 2 words after)

Gold \ Results	Confirmed	Negated	Speculative	Pending	Patient Ant.	Familiar Ant.	na	Recall
Confirmed	387	14	0	1	4	0	2	0,949
Negated	30	171	0	0	0	0	0	0,851
Speculative	14	0	0	0	1	0	0	0,000
Pending	12	0	0	1	0	0	0	0,077
Patient Ant.	42	3	0	0	3	0	1	0,061
Familiar Ant.	11	0	0	0	0	1	0	0,083
na	36	1	0	0	1	0	0	0,000
Precision	0,727	0,905	0,000	0,500	0,333	1,000	0,000	

Total Accuracy: 0.765

matrix corresponds to the true positives for each class, the last row to the precision and the last column to its recall measure.

Regarding the results illustrated in the confusion matrix, our model has a better performance using a nine-word context. This fact rejects the hypothesis that having the entity within the n-gram would improve the performance of the classification. Additionally, as it can be seen, improved results are related to the most frequent classes: Confirmed and Negated. In contrast, we do not have conclusive results with the other classes.

It is possible that the marginal improvement seen in our results is a consequence of having, on the one hand, an unbalanced training set, which includes far more cases of the most common types of senses in which the entities are included in the EHR (confirmed or negated). On the other hand, our assumption regarding the impact of quality (misspelling in particular) did not prove as problematic in the i2b2 test set; however, we do expect other EHR contexts (where entry and coding practices differ) to have more significant quality issues. We did manage to oversample the test to deal with the unbalanced corpus, but the improvement was insignificant. Additional techniques to deal with our unbalanced set can be tested (e.g. Bingel & Haider, 2014). Furthermore, despite the minimal added value of including fuzzy and stemmed NER, our general approach allows for additional NER techniques to be included in a similar fashion.

CONCLUSION

This paper presents an approach that combines named entity recognition and concept attribute labeling over EHR narrative texts. The objective is to recognize important entities in health records such as procedures, medical labs, diagnosis, as well as the context where these entities are written, in order to determine whether they are confirmed, negated, speculative, pending or an antecedent.

Named entity recognition in the healthcare domain is difficult because of the particular and unique characteristics of texts in health records (codified, condensed, jerky language). Due to this, we combine a direct match technique with fuzzy and stemmed matching during the recognition phase. This phase was tested using the i2b2 NLP Data Set which includes a gold standard with annotations. Experimental results on this data set shows an improvement of the recall while having a limited impact on precision.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nevertheless, when analyzing the text, the proposed methods do not take into account the surrounding words (the context) appearing near a named entity candidate.

The second phase involves the analysis of the context of the recognized entity using a concept attribute labeling method that includes six different classes: Confirmed, negated, speculative, pending, patient antecedent, familiar antecedent. This phase comprises an ensemble classification method including SVM and neural networks models. Experimental results over this phase are promising, especially when the entity is confirmed or negated. The results over the other classes are not conclusive yet, mainly because the training and test set do not include a sufficient set of cases.

A limitation of this study is the use of an unbalanced training set, resulting in equally unbalanced assignation by the NER algorithms. That suggests future work using techniques oriented to deal with unbalanced training sets in machine learning. In addition, we are going to improve the concept attribute labeling using the punctuation and part of speech tags as input for the classification model. Our research also suggests that it is important to include in a future version of our solution an engine for regular expression annotations in order to increase the precision and recall of the non-prevalent classes. We made preliminary tests using regular expressions patterns for the speculative class and obtained promissory results; now we have to extend it for the other infrequent classes. Finally, the ensemble classification is going to be enriched with other methods, including k-nearest neighbors, decision trees and deep learning.

REFERENCES

Aramaki, E., Miura, Y., Tonoike, M., Ohkuma, T., Mashuichi, H., & Ohe, K. (2009, June). Text2table: Medical text summarization system based on named entity recognition and modality identification. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing* (pp. 185-192). Association for Computational Linguistics. 10.3115/1572364.1572390

Bingel, J., & Haider, T. (2014). Named-Entity Tagging a Very Large Unbalanced Corpus. Training and Evaluating NE classifiers. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation LREC '14* (pp. 2578-2583).

Casillas, A., Pérez, A., Oronoz, M., Gojenola, K., & Santiso, S. (2016). Learning to extract adverse drug reaction events from electronic health records in Spanish. *Expert Systems with Applications*, *61*, 235–245. doi:10.1016/j.eswa.2016.05.034

Chang, J. T., Schütze, H., & Altman, R. B. (2002). Creating an online dictionary of abbreviations from MEDLINE. *Journal of the American Medical Informatics Association*, *9*(6), 612–620. doi:10.1197/jamia.M1139 PMID:12386112

Dong, X., Qian, L., Guan, Y., Huang, L., Yu, Q., & Yang, J. (2016, August). A multiclass classification method based on deep learning for named entity recognition in electronic medical records. In Proceedings of the Scientific Data Summit (NYSDS), New York (pp. 1-10). IEEE.

Henriksson, A., Dalianis, H., & Kowalski, S. (2014, November). Generating features for named entity recognition by learning prototypes in semantic space: The case of de-identifying health records. In *Proceedings of the 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 450-457). IEEE.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Henriksson, A., Kvist, M., Dalianis, H., & Duneld, M. (2015). Identifying adverse drug event information in clinical notes with distributional semantic representations of context. *Journal of Biomedical Informatics*, *57*, 333–349. doi:10.1016/j.jbi.2015.08.013 PMID:26291578

Jiang, M., Chen, Y., Liu, M., Rosenbloom, S. T., Mani, S., Denny, J. C., & Xu, H. (2011). A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries. *Journal of the American Medical Informatics Association*, *18*(5), 601–606. doi:10.1136/amiajnl-2011-000163 PMID:21508414

Jimeno, A., Jimenez-Ruiz, E., Lee, V., Gaudan, S., Berlanga, R., & Rebholz-Schuhmann, D. (2008, April). Assessment of disease named entity recognition on a corpus of annotated sentences. BMC bioinformatics, 9(Suppl. 3). doi:10.1186/1471-2105-9-S3-S3

Leaman, R., Khare, R., & Lu, Z. (2015). Challenges in clinical natural language processing for automated disorder normalization. *Journal of Biomedical Informatics*, *57*, 28–37. doi:10.1016/j.jbi.2015.07.010 PMID:26187250

Levenshtein, V. I. (1966, February). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics, Doklady*, *10*(8), 707–710.

Lindberg, D. A., Humphreys, B. L., & McCray, A. T. (1993). The unified medical language system. In *IMIA Yearbook* (pp. 41-51).

Menasalvas, E., Rodriguez-Gonzalez, A., Costumero, R., Ambit, H., & Gonzalo, C. (2016, October). Clinical Narrative Analytics Challenges. In *Proceedings of the International Joint Conference on Rough Sets* (pp. 23-32). Springer International Publishing. 10.1007/978-3-319-47160-0_2

Paleo, B. W. (2007, August). An approximate gazetteer for GATE based on levenshtein distance. In *Proceedings of the Twelfth ESSLLI Student Session* (p. 197).

Quimbaya, A. P., Múnera, A. S., Rivera, R. A. G., Rodríguez, J. C. D., Velandia, O. M. M., Peña, A. A. G., & Labbé, C. (2016). Named Entity Recognition Over Electronic Health Records Through a Combined Dictionary-based Approach. *Procedia Computer Science*, *100*, 55–61. doi:10.1016/j.procs.2016.09.123

Skeppstedt, M., Kvist, M., Nilsson, G. H., & Dalianis, H. (2014). Automatic recognition of disorders, findings, pharmaceuticals and body structures from clinical text: An annotation and machine learning study. *Journal of Biomedical Informatics*, *49*, 148–158. doi:10.1016/j.jbi.2014.01.012 PMID:24508177

Stubbs, A., Kotfila, C., Xu, H., & Uzuner, Ö. (2015). Identifying risk factors for heart disease over time: Overview of 2014 i2b2/UTHealth shared task Track 2. *Journal of Biomedical Informatics*, *58*, S67–S77. doi:10.1016/j.jbi.2015.07.001 PMID:26210362

Stubbs, A., & Uzuner, Ö. (2015). Annotating risk factors for heart disease in clinical narratives for diabetic patients. *Journal of Biomedical Informatics*, *58*, S78–S91. doi:10.1016/j.jbi.2015.05.009 PMID:26004790

Uzuner, Ö., South, B. R., Shen, S., & DuVall, S. L. (2011). 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association*, *18*(5), 552–556. doi:10.1136/amiajnl-2011-000203 PMID:21685143

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Xu, Y., Tsujii, J., Eric, I., & Chang, C. (2012). Named entity recognition of follow-up and time information in 20 000 radiology reports. *Journal of the American Medical Informatics Association*, *19*(5), 792–799. doi:10.1136/amiajnl-2012-000812 PMID:22771530

Zhang, S., & Elhadad, N. (2013). Unsupervised biomedical named entity recognition: Experiments with clinical and biological texts. *Journal of Biomedical Informatics*, *46*(6), 1088–1098. doi:10.1016/j.jbi.2013.08.004 PMID:23954592

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 7(1); edited by Anastasius Moumtzoglou; pages 1-15; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 18
Knowledge Management Process-Oriented Strategy for Healthcare Organizations

Nurhidayah Bahar
UCSI University, Malaysia

Shamshul Bahri
University of Malaya, Malaysia

ABSTRACT

This article explores Knowledge Management (KM) practices among doctors and nurses in Malaysia. A total of 59 interviews were conducted with doctors and nurses from two hospitals. The data analysis employed in vivo coding and process coding techniques. The findings suggest a process-oriented strategy for managing knowledge among doctors and nurses in a clinical work environment. The development of this strategy can help the healthcare workers and management to evaluate and further improve their current KM practices. Additionally, this article adds another KM strategy to the literature that is tailored to supporting healthcare organization. Future studies may want to replicate the proposed strategy in different settings such as other clinical or non-clinical departments within the hospital, other public or teaching hospitals or private hospitals.

INTRODUCTION

Knowledge plays an important role in today's organizations. It facilitates decision-making capabilities and builds a learning organization (Garvin, 2003). The collection, creation, and application of knowledge have become critical factors in achieving organizational competitiveness and increasing organizational performance (Asrar-ul-Haq & Anwar, 2016; Supyuenyong, & Swierczek, 2013). Consequently, Knowledge Management (KM) is crucial for organizations that wish to promote best practices, increase their chances of success, and create new business knowledge (Adams & Lamont, 2003; Chapman & Magnus-

DOI: 10.4018/978-1-7998-1204-3.ch018

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

son, 2006). KM refers to the practice of selectively applying knowledge from previous experiences of decision making to current and future decision-making activities with the express purpose of improving organizational effectiveness (Jennex, 2005).

Healthcare has embraced KM. Healthcare aims to enhance the overall quality of lives, meeting the health needs of target populations. This aim is achieved through clinical activities such as assessment, diagnosis, treatment, monitoring, and prognosis. These activities involve knowledge seeking and exchange among clinicians which assist in clinical decision-making. Nevertheless, KM does not yet exist as a full-fledged solution in many healthcare organizations. Few studies have explored knowledge management in the context of healthcare, especially in terms of the contributions of healthcare practitioners in developing systematic KM process in this sector (Abidi, 2001; Agarwal et al., 2011; Beveren, 2003).

This study was motivated by certain weaknesses in previous studies on KM in the healthcare setting. They often focus on a single step of the process such as knowledge acquisition, retention, or dissemination. Few studies have considered the entire KM process from start to finish (Wills et al., 2010). Therefore, these studies were unable to identify how each step affects the next step of the KM process and how it is affected by the previous one. Additionally, many previous studies separated analysis of the processes, people, and technologies employed in the management of knowledge in a healthcare setting (Bhatt, 2000). As a result, these studies missed many nuances that can affect the successful implementation of KM strategy in healthcare. Thus, there is a need to study processes, people, and technologies in tandem.

This study investigates how and when KM occurs in healthcare organizations and the activities required for handling knowledge among doctors and nurses. The study aims to determine the activities needed for managing knowledge and the steps of the KM process, as well as develop a KM process model for healthcare organizations as part of a strategy for implementing systematic KM in this sector.

This study will benefit both healthcare researchers and practitioners. For healthcare researchers, this study contributes to understanding of how knowledge is managed in the healthcare setting. For healthcare practitioners, the findings from this study will lead them to the most appropriate strategies for managing knowledge in the healthcare setting. Hopefully, this knowledge will enhance the success rate of KM implementation in healthcare.

The rest of this paper is structured as follows. The following section presents a review of the relevant literatures followed by methodology, findings, and discussion. The final section discusses study contributions and future research agenda.

LITERATURE REVIEW

KM Strategy as Part of KM Success Model

A KM strategy is an important construct in many KM Success models. KM Success model is useful in predicting success of the design and implementation of a KM initiative (Jennex, 2017). A KM strategy specifically focuses on identifying knowledge users, knowledge for capture and reuse, and process planning, formatting, and context of the knowledge to be stored (Jennex, 2012).

In 1998, Jennex et al. (1998) proposed a KM success model after an ethnographic case study of KM in an engineering organization. They did this by adopting the DeLone and McLean IS Success model (1992). They customized the dimensions to reflect the System Quality and Use constructs relevant in an Organizational Memory System (OMS). The model was then modified by Jennex and Olfman (2002)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

after conducted a longitudinal study in an engineering organization. Their model was also based on the DeLone and McLean (2003) revised IS Success model, and expands on the OMS Success model by including the Information Quality construct.

Another KM Success model was developed by Holsapple and Joshi (2001). This model reflects KM success is based on understanding a process-oriented KM strategy. Subsequent to this study, Massey, Montoya-Weiss and O'Driscoll (2002) develop a process-based KM Success model from their Nortel case study. Their finding suggests that KM cannot be applied generically and that a process approach to KM will assist organization on KM application to improve its performance.

Subsequent to this model, Jennex and Olfman (2006) proposed a model incorporating their experience to design a Knowledge Management System (KMS). Additionally, they incorporated other KM/KMS success factor research from the literature. The model consists of three dimensions: System Quality, Knowledge Quality, and Service Quality. Each dimension consists of three constructs. The KM strategy/process is one of the constructs of Knowledge Quality.

In 2017, Jennex reexamined the Jennex Olfman KM Success Model by incorporating the revised knowledge pyramid. Quantitative research into the artefacts of KM success and targeted papers have used the model. This model consists of System Quality, which reflects how well the KMS assists user in capturing, finding, retrieving, manipulating, and using the knowledge; Knowledge Quality, which reflects the usefulness and accuracy of the content as well as its ability to aid users' duties; and Service Quality, which reflects an organization's ability in ensuring that the KMS provides the expected benefits from knowledge use. KM strategy is one of the constructs in Service Quality that addresses identifying KM goals, alignment, metrics, and knowledge sharing/use incentives.

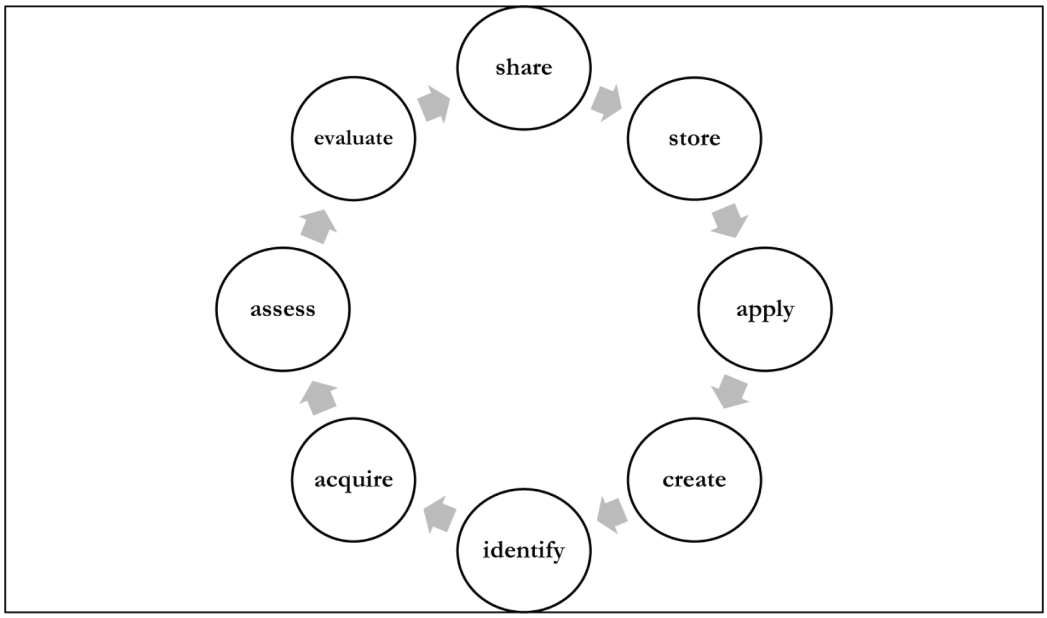
It is evident that KM strategy is vital to successful KM. Many scholars have incorporated KM strategy into their KM/KMS Success models (Massey et al. 2002; Maier, 2002; Jennex & Olfman, 2006). Realizing the importance of KM strategy, Jennex (2012) identified important KM strategy components by examining at successful KM projects/initiatives. He considers the components of a KM strategy as follows: (1) creation/modification of knowledge associated to key performance indicators; (2) increase the awareness/mapping of knowledge sources and users; (3) creation of new/additional knowledge capture processes; (4) changes to the way an organization assesses knowledge use; (5) increase resources for KM systems and its repositories; (6) changes to an organization' s KM goals; and (7) changes in the organization's incentives for knowledge use and sharing. The study found that successful KM project and initiatives are more inclined to the items in the list. Additionally, the findings show that the KM strategy items proposed in the literature are supported by real-world KM project and initiatives.

KM Process Model

In order to identify the activities required for handling knowledge in healthcare organizations, the researchers examined the existing process models. Forty KM process models from the literature were identified and examined (see Table 3 in Appendix 1). Within these models, 179 distinctive terms represent KM activities/process/stages were identified. Some terms more frequently used in various literatures include 'Share/Sharing' within 15 models, 'Create/Creation' in 13 models, 'Apply/Application' in 10 models and 'Storing/Store' in 7 models. These various terms have been classified and grouped into eight broad categories of KM activities (see Table 4 in Appendix 2). The categories are as follows: (1) share; (2) store; (3) apply; (4) create; (5) identify; (6) acquire; (7) assess, and (8) evaluate. Accordingly, the generic activities in most KM models is shown in Figure 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Generic knowledge management model



The notion of cyclical sequence of knowledge processing steps demonstrates that knowledge management is not a simple sequential process (McAdam & McCreedy, 1999; Bukowitz & Williams, 2000).

Knowledge Management in Healthcare Setting

Many KM studies conducted in healthcare organizations focus on decision-support services, examining the role of ICT in facilitating KM practice, assessing healthcare workers' role and their practice in managing knowledge, and developing conceptual models (Abidi, 2001; Bhargava, 2013; Sullivan et al., 2015). For example, Abidi (2001) introduced Strategic Healthcare Decision-Support Services (SHDS) to improve the delivery of quality healthcare services. SHDS acts as a suite of knowledge or data-driven, strategic and decision-support services derived from both healthcare data and health enterprise knowledge bases. In addition to KMS, scholars have also investigated the roles of other Information Communication Technologies (ICT) such as technical infrastructure, decision support architecture, tools for knowledge sharing (Whiddett et al., 2012) and enhancing knowledge transfer (Sheng et al., 2013).

Scholars are also concerned in assessing healthcare workers' roles and behaviors. For instance, Dehaghi et al. (2015) conducted a study to investigate the relationship between the quality of work life of nurse managers and their participation in KM implementation. They found that the participation in implementing KM among nurse managers may increase if there is an improvement of nurse managers' work life quality, such as in decision-making. Finally, studies relating to KM in healthcare have considered model development. Beveren (2003) explored KM within an Australian regional healthcare organization and found a number of obstacles inherent in its organizational structure. He concluded that new models, tools, and techniques for KM specific to the health sector are necessary. Subsequent to this finding, Lau (2004) described the conceptual organizing scheme for managing knowledge within the healthcare. This

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

focuses on the concepts of production, use, and refinement of three types of knowledge sources, namely policy, evidence, and experience. According to Lau (2004), these concepts can be operationalized through a set of KM methods and tools tailored to the healthcare setting.

RESEARCH METHOD

The study was divided into two stages. First was the development of the generic process model based on examination of existing KM process models from the literature. Jennex (2005) proposed a model that is based on other widely supported models and is more generalizable. The generic process model was used in developing the interview questions. To generate the model, the researchers searched knowledge management related journals using three key phrases: knowledge management model, knowledge management process and knowledge framework. Forty KM process models from the literature were examined and a generic process model was developed. The generic model consists of eight activities: (1) share; (2) store; (3) apply; (4) create; (5) identify; (6) acquire; (7) assess; and (8) evaluate.

Second, the study included interpretive case studies in order to gain in-depth understanding and insight into the KM practice in clinical care environment. This study was conducted in two hospitals: a public hospital (Case I) and a teaching hospital (Case II). The healthcare organizations used for this case study are among the largest. They are located in central Malaysia that serve a large population ranging from around 680,000 to 780,000 people. These organizations are responsible for improving the health of the public. Primary healthcare is provided in the community and includes general practitioners, psychiatrists, dentists, pharmacists and wide range of health professionals. Input for the case studies was received from doctors and nurses from the generic medical department in both hospitals.

The researchers were required to seek approval from the Ministry of Health's and hospital's Medical Ethics Committee before conducting the field study. Prior to ethical clearance, a co-supervisor from among the doctors of both hospitals was appointed and approval was sought from each hospital's Medical Ethics Committee through an online application. Upon completion of the field study, the researchers were required to report the findings to the co-supervisors and produce the final report.

The researchers followed a purposeful sampling approach, gathering informants to collect responses to the formulated research questions. Interviews were conducted with twenty-nine doctors and thirty nurses from the generic medical department in both hospitals. All informants first received information about the study and the interview via e-mail, after which interviews were scheduled either via e-mail or phone. Topics covered during the interviews included the backgrounds and roles of interviewees, their beliefs structures about the nature of valuable knowledge, knowledge search and/or acquisition patterns and modes of using knowledge.

The different types of roles selected in this study were treated as multiple sources for assessing similarities and differences in KM practices. Doctors and nurses are most actively engaged in direct patient care and provide leadership in the clinical work environment. They work closely together to make collective decisions for patient care. They collaborate by sharing personal and professional knowledge and joint responsibility to ensure the utilization of best knowledge to produce positive patient outcomes.

In addition to providing patient care, doctors and nurses play a leading role in generating medical evidence and engaging in clinical research. They collect, track, and analyze clinical and patient data to help build the scientific groundwork for clinical practice, avoidance, and enhanced patient health outcome, as well as evaluating and enhancing their own practice. Table 1 provides a breakdown of the informants.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Informants

Case	Doctor	Nurse
Case I	17	18
Case II	12	12
Total	**29**	**30**
Grand Total	**59**	

The informants had from three to thirty years' work experience. The interviews were conducted at informants' premises and lasted about forty-five minutes to one and a half hours. To ensure the privacy of the research informants and to comply with regulations, the researchers followed the approval and notification procedures as prescribed by the hospital's Medical Ethics Committee. The informants were informed about the research and its purposes. An information sheet was provided regarding their rights to confidentiality and anonymity. Their consent was also obtained prior to participation. The interviews were conducted at informants' premises fir forty-five minutes to one and a half hours. They were electronically recorded with the consent of the parties involved for data reorganization and analysis purposes.

Data Analysis

Data analysis was carried out as the data collection progressed. Interweaving data collection and analysis from the very start helps to cycle back and forth between thinking about the existing data and generating strategies for collecting new data (Miles, Huberman & Saldana, 2013). The interviews were recorded, while notes were taken to identify the major points. The recordings were then transcribed verbatim.

First, cycle coding was used to initially summarize segments of data. The researchers employed in vivo coding and process coding techniques. In vivo coding uses words or short phrases from the participant's own language in the data record as codes, while process coding method uses gerunds ("-ing" words) to connote observable and conceptual action in the data (Miles, Huberman & Saldana, 2013). Table 2 shows samples from first cycle coding methods.

FINDINGS

Analysis indicates that doctors and nurses from the two hospitals have similar patterns in terms of clinical practices. These practices are examples of decision-making and knowledge utilization situations in healthcare organizations. Based on their practices, a KM process-oriented strategy for clinical work environment was developed. Figure 2 depicts the process model and the linkages among the activities. Building on the generic KM model, KM practice among the doctors and nurses consists of ten activities: (1) identifying; (2) discovering; (3) acquiring; (4) gathering; (5) analyzing; (6) capturing; (7) organizing; (8) disseminating/communicating; (9) applying/experimenting; and (10) evaluating:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Informants quotes underlying first cycle coding

Informant Quotes	First Cycle Coding	
	In Vivo Coding	**Process Coding**
• Sometimes we encounter situation where patients suffering from rare diseases. It requires us to search information worldwide. For example, some foreign patients were infected by a disease that origin from their country. It usually takes a lot of work to diagnose by identifying and seeking information from various external sources.	We know what we don't know.	Identifying
• The changing nature of knowledge requires us to bring our know-how to date. It is therefore very important to search for the latest information especially the diagnosis and prognosis made by the doctors. This is very important that places a patient between life and death.	We know how to find what we don't know.	Discovering
• Normally, we refer to CPG in conducting clinical care. However, CPG and internal resources alone are insufficient. We do consult with expert from other hospitals to assist decision-making and diagnosis. • I access online resources such as UpToDate website to get the latest information.	We know where to find what we don't know.	Acquiring
• Nurse will perform the first-level assessment and provide us (doctors) the outcome. • I will verify the outcomes from the assessment before making decision. • Patient's story is important because they know their body better. • Lab tests and medical record can be accessed from our electronic medical record system. It aids the process of deciding what to do.	We make judgments based on numerous evidences.	Gathering
• Based on patient's data, I have to deliberate on what could happen and thus look for evidence to support my assumption. • I will have to carefully review patient's records and examine lab test results. • We analyze problems to figure out solutions.	We look for patterns and compare with precious cases.	Analyzing

- **Identifying Knowledge/Information:** In the first activity, doctors, and nurses indicate what knowledge/information is involved in order to perform their clinical care process such as assessment, diagnosis, treatment, monitoring, prognosis, and patient disposition. This study found four types of knowledge/information in the clinical work environment; (1) personal knowledge and competency; (2) patients' experience of illness and health condition; (3) clinical evidence and practice guideline; and (4) technological know-how. As both doctors and nurses require strong knowledge and competence to care for their patients, they must keep their knowledge and skills up-to-date by engaging in lifelong learning. According to one of the doctors:

It is important to keep our knowledge and skills up-to-date from time to time. By attending seminar, training, and medical course, we (doctors) will get the latest updates on guidelines and procedures to reflect the latest disease symptoms.

- **Discovering Knowledge/Information:** Once the required knowledge is identified, the next step is discovery. In this context, discovery means finding and locating the required knowledge/information. This study found common techniques practiced by the doctors and nurses in searching for personal knowledge, patient experience, clinical evidence and guideline as well as the know how to use IT/IS in their profession. Doctors and nurses discover or update knowledge through medical school, seminars, continuing medical/nurse education, research and development, clinical practice and online medical reference sites. One of the doctors quoted:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The changing nature of knowledge requires us to bring our know-how to date. It is very important to search for the latest information especially the diagnosis made by doctors will place a patient between life and death.

- **Acquiring Knowledge:** Discovery activities will lead to knowledge acquisition, in which doctors and nurses extract information from the identified sources to enhance their knowledge. They validate whether the extracted information is what they need. If there is missing information, the process will go back to identifying then discovering knowledge. Acquiring knowledge from the identified sources by doctors and nurses will lead to enhancing personal knowledge, skill, and competency; better understanding on patient's health condition; ability to manage different clinical problems and situations; and additional evidence to support judgments and decisions. As they gain more experience over time, they will be able to build a set of cases with identifiable patterns that will serve as valuable background knowledge when dealing with their day-to-day situations. These activities are exemplified by the following observations:

In straightforward or common situations, I often make such decisions informally; For example, during a flu epidemic, a healthy adult who has had fever, aches and harsh cough for 2 days is likely to be recognized as another case of influenza. (Doctor)

For a complex situation, we have to make good judgment.... beyond the scope of protocols and guidelines. So, we must be able to recognize when to apply protocols.... recognize changing circumstances and adapt them to the situation at hand. (Doctor)

- **Gathering Knowledge/Information:** The information use in clinical care process is gathered through direct observations, examinations, patient-related history their concerns, EMR or hard-copy reports, and clinical lab test results. Nurses take charge to collect patient information and medical history, perform initial assessments, and order clinical lab tests such as blood, urine and swab tests. This information will then be presented to doctors for further assessment and judgments. At this stage, doctors and nurses will also gather opinions from peers and seniors. Additionally, doctors will look for evidence from best practice guidelines such as Clinical Practice Guidelines (CPGs), Quick Reference (QR), and Training Manual (TM). These guidelines consist of detail instructions of the following situations: (1) determining the appropriate screening tests or diagnostics; (2) specifying medical or surgical services; (3) the duration of stay in hospital; or (4) other details of clinical practice. One of the nurses quoted:

I normally start with talking to the patient before assessment. Patient's story is important because they know their body better. On top of that, I accessed lab test results and medical record from our electronic medical record system. It aids the process of deciding what to do.

- **Analyzing the Gathered Knowledge:** Once knowledge and information are acquired, doctors and nurses systematically determine relevant points before analysis. This involves reviewing and examining data. Based on the analyzed data, doctors make judgments on what they think could be happening, looking for patterns and assessing if there is any additional information needed. At

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

this stage, doctors play an important role in the diagnosis of patient. They must have the ability to combine exact and interpreted information with their existing knowledge base in order to make accurate decisions.

- **Capturing Knowledge/Information:** After analysis, useful knowledge and information will be recorded. The outcomes from solving the clinical problems are beneficial for future use and reference. Therefore, doctors and nurses feel that it is important to remember and record how to resolve certain problems. They keep information in a useful form. Such knowledge/information is mainly recorded in individual memory. Details such as patient's information, medical report, physical examination result and final diagnosis are recorded in a hard copy form or electronic medical record. This activity is exemplified in the following quotes:

I know how to conduct some clinical protocols from practical session. For example, to examine patient's abdomen requires a few steps starting from observation and then sequentially performing auscultation, palpation, and percussion. I can do it better because I have seen how my seniors are doing it. (Nurse)

Well, most of the times I have no chance to systematically keep notes from my clinical experience… But, if there's anything new, I am sure it is recorded in my personal memory. It's natural process to me. (Doctor)

- **Organizing the Knowledge:** Once the important knowledge/information is captured, it must be organized accordingly. At this stage, the information that comes from various sources should be restructured. The study found that knowledge in healthcare organization exists in individual experts, databases, documented organizational procedures such as Clinical Practice Guidelines (CPGs) and standard operating procedures (SOPs). Analysis suggests that in the clinical work environment, knowledge is mainly possessed by individuals (i.e. clinicians) as they are the main actors to capture knowledge. This includes collecting and organizing the documents in a meaningful manner for easy access to and retrieval of knowledge content. Each doctor and nurse captures and organizes knowledge in their own way in an individually embedded manner. Knowledge is normally stored in personal notebooks and documents.
- **Disseminating/Communicating Knowledge:** Doctors and nurses spread their intellectual knowledge amongst themselves. There are various tools which doctors and nurses normally use to disseminate information. Informal methods include social media and instant messaging applications, while more formal methods include medical/nurse educational training and courses, meetings, and emails. These methods allow for the dissemination of detailed information and data to others as well as electronic news services from organization websites. Doctors and nurses also utilize social media to create awareness and educate the public by posting health related information or articles. In that way, they reach a much broader audience of people. Most doctors claimed that the organizations encourage them to publish educational articles in medical journals so that they can share their experiences and educate others. After clinical practice, they normally write a report for future reference, especially when they come across a novel issue or finding that they would like to write about. From there, doctors might start writing an educational article.
- **Applying/Experimenting Knowledge:** At the clinical level, knowledge use is seen as a process through which clinicians formulate a solution in order to solve their day-to-day problems in providing patient care. They apply their knowledge to make judgments and prioritize these judgments based on the current patients, situation and environment. This means eventually deciding what

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to do and how to do it, as well as determining who should be informed or consulted. When they apply, and at the same time they experiment where they can evaluate whether the outcome of the decision determined is the desired outcome that has been achieved, as represented in the following descriptions:

Every patient has a different reaction to medicine or treatment. We need to find out the most suitable treatment for them. When the patient comes back to us with better health condition after our treatment, we know that the treatment given is suitable for such cases. (Doctor)

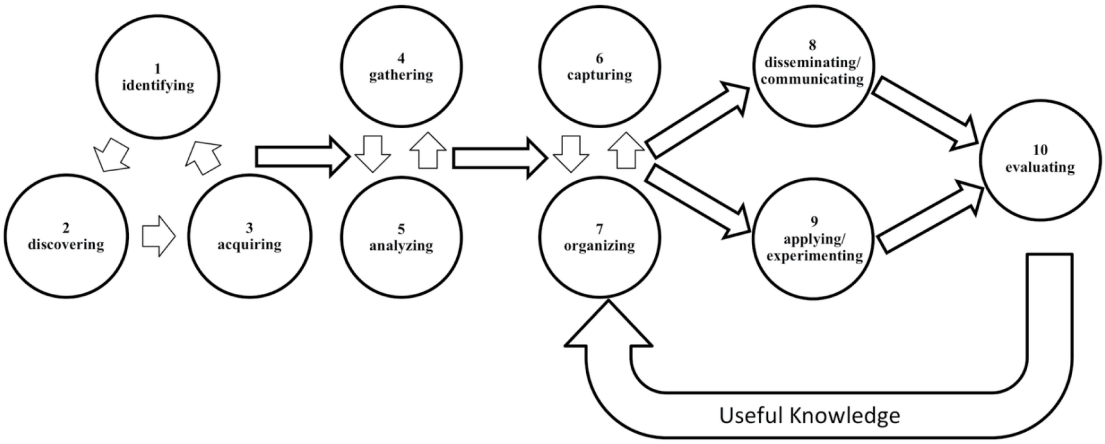
Evaluating an Outcome/Result

Doctors and nurses are active decision makers who determine who needs what and when. Once a decision is made, evaluation must take place to ensure decision effectiveness. Based on the interviews, doctors and nurses practice post-treatment evaluation to assess the significance or quality of outcomes derived from the previous activities i.e. clinical decisions. This activity is essential to clarify whether a decision has been effective and to help to refine their existing knowledge. As part of the evaluation process, doctors are responsible to administer a particular medical treatment to be given to the patients. The idea is to ensure that they're making the best choice and avoid negative consequences. It also helps to enhance practice delivery in future.

DISCUSSION

To recap, the generic KM model that was adopted in this study consists of eight elements: (1) share; (2) store; (3) apply; (4) create; (5) identify; (6) acquire; (7) assess; and (8) evaluate. As the researchers conducted this study in healthcare organizations, five additionally elements were identified: (1) discovering; (2) gathering; (3) analyzing; (4) capturing; and (5) organizing.

Figure 2. Knowledge management process-oriented strategy



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The KM process-oriented strategy proposed knowledge management activities in healthcare begins with identifying, discovering, and acquiring knowledge. Many studies, however, collapse these activities into the category of identifying (Lai & Chu, 2002; Andersen & APQC, 1996; Tikhomirova et al., 2008). In the healthcare setting, these activities are distinct and requires particular action rather than a combination, as each involves particular action from the knowledge seeker.

Few studies were found to have included assessment as part of their KM process model. Assessment activities in previous models tend to occur towards the end of the process, indicating that the models did not suggest knowledge filtering before use and storage (Meyer & Zack, 1996; Karadsheh et al., 2009, Amirkhani et al., 2012; Chauhan et al., 2016; Sullivan et al., 2015). This process is unacceptable in a healthcare setting. Knowledge must be assessed for fitness before use. The doctors and nurses see the importance of assessing the knowledge before passing it to the next stage. The process of acquiring knowledge involves diverse sources, and such knowledge may be inaccurate or worthless (Sun & Hao, 2006).

Minimal studies have included evaluation at the end of the process. In healthcare setting, this activity is crucial to determine the relevance and effectiveness of the decision-making process. For example, if the treatment plan is found to be effective, knowledge will be captured and organized for future reference. However, if the action taken is deemed to be irrelevant or insufficient, they will then review, reflect, and adapt the lessons learned. Doctors and nurses learn from events so that they are not repeated.

KM is not a linear process, and the process-oriented strategy in this study reflects that point. There are three iterative stages in this process; (1) identifying, discovering, and acquiring; (2) gathering and analyzing; and (3) capturing and organizing information. These iterative processes represent a cycle of operations before achieving the desired result or outcome. For example, in order to acquire information in solving clinical problem, doctors will identify the required information and choose the possible sources to search for (discover) the information. Once information has been found to be irrelevant or additional information is required, doctors will repeat previous activities until the desired information meets their needs. Finally, the overall components of a KM strategy proposed in this study is consistent with that of Jennex (2008, 2012), as a KM strategy must include identification of users, sources of knowledge, knowledge to be captured, and how captured knowledge is to be stored and represented, among other factors.

CONCLUSION

This study explores knowledge management in healthcare organizations. The healthcare industry faces diverse factors characterizing patients (e.g. physical traits, experience of illness and medical history), professional disciplines (e.g. doctors, nurses, pharmacists, radiologists and administrators), treatment options, delivery processes, and the interests of various stakeholder groups (patients, regulators, health agencies and civil society organizations) (Fichman et al., 2011; Kothari et al., 2011). Healthcare organizations are typically inter-organizational collaborations with a multitude of actors and stakeholders.

The study found that both healthcare organizations do not have a full-fledged KM in practice. Therefore, this study proposes a process-oriented KM strategy in implementing systematic KM in healthcare. This model has been tailored exclusively to healthcare organizations, with their unique characteristics and limitations. Studies on knowledge management in healthcare setting is lacking in integrated knowledge management process model. Healthcare processes are knowledge-intensive in nature (Myllärniemi et al., 2012). This study has helped to uncover a knowledge management process-oriented strategy that

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

suits clinical processes and decision-making. The proposed strategy aims to achieve greater and more relevant application of KM in the healthcare setting. This will encourage comprehensive knowledge management views, activities, and technologies in healthcare organizations.

The development of this process-oriented strategy will help healthcare workers and management to evaluate current knowledge management practices and potential improvements. Solving problems and making optimal decisions in healthcare is heavily dependent on access to knowledge. Thus, healthcare organizations should provide opportunities to incorporate knowledge management practices provide the best possible healthcare through operational excellence. These goals are achievable if a well-organized and effective strategy for knowledge management in healthcare is implemented.

Based on the current findings, this study indicates the following avenues for further research. Future studies may replicate the proposed KM strategy in different settings within the broader healthcare context. Such settings may include other clinical or non-clinical departments, other public or teaching hospitals, or private hospitals. Future research would also benefit from collecting data from different informant profiles (e.g. pharmacists, radiologists and other hospital staff members).

The proposed process-oriented strategy is focused on clinical activities from only two hospitals. Thus, the results may lack generalisability. However, its use in other contexts should not be dismissed as long as necessary adjustments are made.

REFERENCES

Abidi, S. S. R. (2001). Knowledge Management in healthcare: Towards 'knowledge-driven' decision-support services. *International Journal of Medical Informatics*, *63*(1-2), 5–18. doi:10.1016/S1386-5056(01)00167-8 PMID:11518661

Adams, G. L., & Lamont, B. T. (2003). Knowledge Management systems and developing sustainable competitive advantage. *Journal of Knowledge Management*, *7*(2), 142–154. doi:10.1108/13673270310477342

Agarwal, A., Garg, S., & Pareek, U. (2011). Strengthening human resources practices in healthcare in India, the road ahead. *Journal of Indian Academy of Clinical Medicine*, *12*, 38–43.

Amirkhani, A., Tajmirriahi, J., & Mohammad, M. (2012). Assessing the effectiveness of Knowledge Management in empowering and development of human resources: A case study on Ati Luleh Sepahan Company. *Interdisciplinary Journal of Contemporary Research in Business*, *3*(12), 131–147.

Anderson, A. (1996). *The APQC*. The Knowledge Management Assessment Tool.

Asrar-ul-Haq, M., & Anwar, S. (2016). A systematic review of Knowledge Management and knowledge sharing: Trends, issues, and challenges. *Cogent Business & Management*, *3*(1), 1–7. doi:10.1080/23311975.2015.1127744

Beveren, J. V. (2003). Does health care for Knowledge Management? *Journal of Knowledge Management*, *7*(1), 90–95. doi:10.1108/13673270310463644

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bhargava, A., Mishra, B., Thakur, A., Dogra, V., Loomba, P., & Gupta, S. (2013). Assessment of knowledge, attitude and practices among healthcare workers in a tertiary care hospital on needle stick injury. *International Journal of Health Care Quality Assurance*, *26*(6), 549–558. doi:10.1108/IJHCQA-04-2012-0035 PMID:24003754

Bhatt, D. (2000). EFQM Excellence model and Knowledge Management implications. EFQM Organization.

Bukowitz, W. R., & Williams, R. L. (2000). *The Knowledge Management Fieldbook*. Financial Times/ Prentice Hall.

Chapman, R. L., & Magnusson, M. G. (2006). Continuous innovation, performance and Knowledge Management: An introduction. *Knowledge and Process Management*, *13*(3), 129–131. doi:10.1002/kpm.252

Chauhan, J., & Raksha, K. P. (2016). Refocusing the Concept and Application of Knowledge Management. *Indian Research Journal of Extension Education*, *15*(1), 118–127.

Dehaghi, Z. H., Sheikhtaheri, A., & Dehnavi, F. (2015). Nurse managers' work life quality and their participation in Knowledge Management: A correlational study. *Iranian Red Crescent Medical Journal*, *17*(1). PMID:25763267

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95. doi:10.1287/isre.3.1.60

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30. doi:10.1080/07421222.2003.11045748

Fichman, R. G., Kohli, R., & Krishnan, R. (2011). Editorial Overview-The Role of Information Systems in Healthcare: Current Research and Future Trends. *Information Systems Research*, *22*(3), 419–428. doi:10.1287/isre.1110.0382

Garvin, D. A. (2003). *Learning in action: A guide to putting the learning organization to work*. Harvard Business Review Press.

Hojabri, R., Borousan, E., & Manafi, M. (2012). Impact of Using Telemedicine on Knowledge Management in Healthcare Organizations: A Case Study. *African Journal of Business Management*, *6*(4), 1604.

Holsapple, C. W., & Joshi, K. D. (2001). Organizational knowledge resources. *Decision Support Systems*, *31*(1), 39–54. doi:10.1016/S0167-9236(00)00118-4

Jennex, M. E. (2005). What is Knowledge Management? *International Journal of Knowledge Management*, *1*(4), i–iv.

Jennex, M. E. (2008). Impacts from using knowledge: A longitudinal study from a nuclear power plant. *International Journal of Knowledge Management*, *4*(1), 51–64. doi:10.4018/jkm.2008010105

Jennex, M. E. (2012, August). Identifying the Components of a Knowledge Management Strategy. In *Proceedings of the 18th Americas Conference on Information Systems, AMCIS18*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jennex, M. E. (2017, January). Re-Examining the Jennex Olfman Knowledge Management Success Model. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. 10.24251/HICSS.2017.567

Jennex, M. E., & Olfman, L. (2002, January). Organizational memory/knowledge effects on productivity, a longitudinal study. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (pp. 1029-1038). IEEE.

Jennex, M. E., & Olfman, L. (2006). A Model of Knowledge Management Success. *International Journal of Knowledge Management*, *2*(3), 51–68. doi:10.4018/jkm.2006070104

Jennex, M. E., Olfman, L., Panthawi, P., & Park, Y. T. (1998, January). An organizational memory information systems success model: an extension of DeLone and McLean's I/S success model. In *Proceedings of the Thirty-First Hawaii International Conference on System Sciences* (*Vol. 1*, pp. 157-165). IEEE.

Karadsheh, L., Mansour, E., Alhawari, S., Azar, G., & El-Bathy, N. (2009). A theoretical framework for Knowledge Management process: Towards improving knowledge performance. *Communications of the IBIMA*, *7*, 67–79.

Kothari, A., Hovanec, N., Hastie, R., & Sibbald, S. (2011). Lessons from the business sector for successful knowledge management in health care: A systematic review. *BMC Health Services Research*, *11*(1), 173. doi:10.1186/1472-6963-11-173 PMID:21787403

Lai, H., & Chu, T. H. (2002). Knowledge Management: A review of industrial cases. *Journal of Computer Information Systems*, *42*(5), 26–39.

Lau, F. (2004). Toward a conceptual Knowledge Management framework in health. *Perspectives in Health Information Management*, *1*(8), 1–11. PMID:18066388

Maier, R. (2002). Foundation of Knowledge Management Systems. In *Knowledge Management Systems* (pp. 17–86). Springer Berlin Heidelberg. doi:10.1007/978-3-662-04380-6_4

Massey, A. P., Montoya-Weiss, M. M., & O' Driscoll, T. M. (2002). Knowledge management in pursuit of performance: Insights from Nortel Networks. *Management Information Systems Quarterly*, *26*(3), 269–289. doi:10.2307/4132333

McAdam, R., & McCreedy, S. (1999). A critical review of knowledge management models. *The Learning Organization*, *6*(3), 91–101. doi:10.1108/09696479910270416

Meyer, M. H., & Zack, M. H. (1996). The design and development of information products. *MIT Sloan Management Review*, *37*(3), 43.

Miles, M. B., Huberman, A. M., & Saldana, J. (2013). *Qualitative data analysis: A methods sourcebook*. SAGE Publications, Incorporated.

Myllärniemi, J., Laihonen, H., Karppinen, H., & Seppänen, K. (2012). Knowledge management practices in healthcare services. *Measuring Business Excellence*, *16*(4), 54–65. doi:10.1108/13683041211276447

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sheng, M. L., Chang, S. Y., Teo, T., & Lin, Y. F. (2013). Knowledge barriers, knowledge transfer, and innovation competitive advantage in healthcare settings. *Management Decision*, *51*(3), 461–478. doi:10.1108/00251741311309607

Sullivan, T. M., Limaye, R. J., Mitchell, V., D'Adamo, M., & Baquet, Z. (2015). Leveraging the power of Knowledge Management to transform global health and development. *Global Health, Science and Practice*, *3*(2), 150–162. doi:10.9745/GHSP-D-14-00228 PMID:26085014

Sun, Z., & Hao, G. (2006). HSM: A hierarchical spiral model for Knowledge Management. In *Proceedings of the 2nd International Conference on Information Management and Business (IMB2006)*, Sydney, Australia, February 13-16 (pp. 542-551).

Supyuenyong, V., & Swierczek, F. W. (2013). Knowledge management process and organizational performance in SMEs. In *Dynamic Models for Knowledge-Driven Organizations* (pp. 77–98). IGI Global. doi:10.4018/978-1-4666-2485-6.ch005

Tikhomirova, N., Gritsenko, A., & Pechenkin, A. (2008). University approach to Knowledge Management. *The Journal of Information and Knowledge Management System*, *38*(1), 16–21.

Whiddett, D., Tretiakov, A., & Hunter, I. (2012). The use of information technologies for knowledge sharing by secondary healthcare organisations in New Zealand. *International Journal of Medical Informatics*, *81*(7), 500–506. doi:10.1016/j.ijmedinf.2012.02.011 PMID:22460023

Wiig, K. M. (1994). *Knowledge Management Foundations: Thinking about Thinking-how People and Organizations Represent, Create, and Use Knowledge*. Schema Press, Limited.

Wills, M. J., Sarnikar, S., El-Gayar, O. F., & Deokar, A. V. (2010). Information systems and healthcare XXXIV: Clinical Knowledge Management systems - Literature Review and Research Issues for Information Systems. *Communications of the Association for Information Systems*, *26*(26), 565–598.

This research was previously published in the International Journal of Knowledge Management (IJKM), 13(4); edited by Murray E. Jennex; pages 73-89; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Table 3. Knowledge management process model

No.	Author	KM Process
1	Wiig (1993)	Build (obtaining, analyzing, reconstructing/synthesizing, codifying and organizing), hold (remembering, accumulating and embedding), pool (coordinating, assembling, accessing and retrieving at group level) and apply.
2	Slater and Narver (1995)	Acquisition of information, information dissemination, shared interpretation
3	Arthur Andersen and APQC (1996)	Create, identify, collect, adapt, organize, apply and share
4	Demerest (1997)	Knowledge construction, knowledge embodiment, knowledge use and knowledge dissemination.
5	van der Spek and Spijkervet (1997)	Conceptualize, reflect, act, and retrospect
6	Meyer and Zack (1999)	Acquisition, refinement, storage/retrieval, distribution and presentation/use.
7	Bukowitz and Williams (1999)	Get, assess, build/sustain, contribute and learn.
8	Crossan et al. (1999)	Intuiting, interpreting, integrating, institutionalizing
9	Probst, Raub and Romhardt (1999)	Knowledge identification, knowledge sharing or knowledge distribution.
10	Boisot (1999)	Codification, abstraction and diffusion
11	Bukowitz and Williams (2000)	Create/capture of knowledge, storage/retrieve, access, distribution, sustain and disposal.
12	Alavi and Leidner (2001)	Creation, storage/retrieval, transfer and application.
13	Gamble and Blackwell (2001)	Observe, gather, hypothesize, contextualize, categorize, map, share, disseminate and simulate.
14	Abou-Zeid, (2002)	Knowledge identification, knowledge generation, knowledge elaboration, knowledge preservation, knowledge mobilization, knowledge presentation and knowledge evaluation.
15	Lai and Chu (2000)	Initiation, generation, modeling, repository, distribution & transfer, use and retrospect.
16	McElroy (2003)	Knowledge claim (validation) and knowledge integration (sharing and disseminating).
17	Frid (2003)	Knowledge chaotic, knowledge aware, knowledge focused, knowledge managed, and knowledge centric.
18	Arostegui (2004)	Capturing, elaborating, transferring, storing and sharing knowledge.
19	Diakoulakis et al. (2004)	Exploration of the external environment, internal scanning, sharing/access of knowledge, retention/systemization of knowledge, combination/creation of knowledge.
20	Lettieri, Borga and Savoldelli (2004)	Acquisition, codification, storage, retrieval, creation, application and diffusion & presentation.
21	Dalkir (2005)	Create/capture, assess, share/disseminate, contextualize, apply/use and update
22	Lee et al. (2005)	Create/capture of knowledge, update, application, utilization
23	Grant (2005)	Knowledge identification, knowledge measurement, knowledge storage & organization, knowledge replication, knowledge integration and knowledge sharing.
24	Baptista et al. (2006)	Capturing, storing, sharing and distributing knowledge
25	Hsia, Lin, Wu and Tsai (2006)	Knowledge creation, knowledge codification, knowledge transfer, knowledge application
26	Sheffield (2008)	Creation, normalization and application of knowledge
27	Botha et al (2008)	Knowledge creation and sensing, knowledge organizing and capture and knowledge sharing and dissemination.

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Continued

No.	Author	KM Process
28	Tikhomirova et al. (2008)	Identification and capture, creation, classification and storage, circulation and distribution and application of knowledge
29	Karadsheh et al. (2009)	Knowledge combination, knowledge evaluation, knowledge filtering (organization, classification and categorization), knowledge repository, knowledge sharing, knowledge application and knowledge performance.
30	Heisig (2009)	Use, identify, create, acquire, share and store.
31	Fugate et al. (2009)	Generation, dissemination, sharing and interpretation of knowledge
32	Huang and Shih (2009)	Creation, storage, distribution and utilization of knowledge
33	Nag and Gioia (2012)	Executive knowledge schemes, executive scanning and knowledge use.
34	Amirkhani et al. (2012)	Specifying strategic knowledge goals, acquiring the required knowledge, assessment and organization of knowledge, sharing knowledge, empowerment and sustainable human resources development
35	Pawlowski and Bick (2012)	Knowledge identification, knowledge acquisition, knowledge development, knowledge distribution/sharing, knowledge preservation and knowledge use.
36	Tuamsuk et al. (2013)	Knowledge identification, creation and acquisition, knowledge storing, knowledge distribution, knowledge application
37	Evans, Dalkir and Bidian (2015)	Identify, store, share, use, learn, improve, and create
38	García-Fernández (2015)	Creation, transfer and storage, and implementation and use
39	Chauhan, Raksha and Pradhan (2015)	Detect and discover (gather, observe and deconstruct) and organize and assess (categorize, contextualize and map).
40	Sullivan et al. (2015)	Knowledge assessment, generation, capture, synthesis, and sharing

Table 4. Seven most frequently discussed groups of knowledge management activities

No.	KM Activities	Total Count in Frameworks
1	Share (15), Share/Disseminate (2), Distribution/Distributing (7), Circulation & Distribution (1), Distribution/ Sharing (1), Transfer/ Transferring (5), Dissemination (5), Contribute (1), Mobilization (1)	38
2	Store (7), Storage/Retrieval (4), Storage & Organization (1), Classification & Storage (2), Organize (1), Organizing & Capturing (1), Organize & Assess (1), Build/Sustain (1), Sustain (1), Capture (4), Capturing (2), Codification (3), Access (2), Retention/Systemization (1), Sustainable (1), Preservation (2), Replication (1), Repository (2), Hold (1)	38
3	Apply (10), Use (8), Apply/Use (1), Utilization (2), Implementation (1), Presentation (3), Adapt (1), Act (1), Simulate (1)	28
4	Create (13), Combination/Creation (1), Create/Capture (1), Creation & Sending (1), Combination (1), Creation & Acquisition (1), Knowledge Construction (1), Generation (4), Development (1), Build (1)	25
5	Assess (3), Assessment and Organization (1), Filtering (1), Validation (1), Interpreting (2), Elaboration (2), Synthesis (1), Evaluate (2), Learn (2), Reflect (1), Retrospect (2), Refinement (1), Performance (1), Measurement (1), Improve (1)	22
6	Identify (8), Identification & Capture (1), Specifying (1), Categorize (1), Map (1), Exploration (2), Knowledge Aware (1)	15
7	Acquire (4), Collect (1), Get (1), Acquire (2), Gather (1), Detect & Discover (1), Scanning (2), Observe (1)	13

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 19
Use-Case Driven Approach for a Pragmatic Implementation of Interoperability in eHealth

Karima Bourquard
IHE – Europe, France

Alexander Berler
IHE – Europe, Belgium

ABSTRACT

Innovation in IT solutions introduces new usages in ehealth that allows the patient to be more active in his (her) health care and well-being. Strong expectations on electronic exchanges or shared medical documents from healthcare professionals and patients are pushing decision makers and stakeholders to develop quickly efficient and structured data enabled health IT infrastructures. One of the challenges of this paradigm shift is to provide at the point of care interoperable solutions in full compliance with regulations, ethics, organizational and technical states of the art. This paper describes this approach developed as a four steps process and defines the basic concepts for designing an interoperability infrastructure. For each of the steps, examples will be provided and discussed accordingly with the different interoperability layers (legal, policy, process, information and technical) developed in the Refined eHealth European Interoperability Framework (ReEIF) (eHealth Network, 2015).

INTRODUCTION

A pragmatic approach for a more efficient deployment of solutions in eHealth is now more and more expected by the user stakeholders as well as by public authorities due to the acceleration of the technological and digital divide and by the paradigm shift in healthcare in term of users:

- Patients develop more interest in their health and well-being and therefore demand a more comprehensive and direct access to their data and care plan;

DOI: 10.4018/978-1-7998-1204-3.ch019

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Innovative solutions provide more information that needs to be filtered depending on their usages and provide adequate access to entrusted users;
- New medical processes have not yet sufficiently evaluated (i.e., behavioral health, precision medicine, etc.).

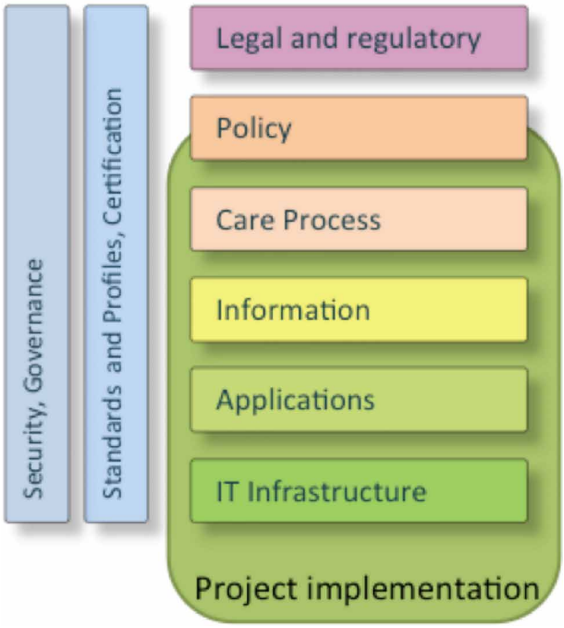
All these new processes require more direct and secure communication channels between healthcare professionals and patients taking into account the need for improved quality of data, extended care pathways where professionals are exchanging information within a collaborative process.

To facilitate the integration of those innovative solutions with the healthcare processes (current and future), and to involve end-users in the structured definition of their needs, a use case driven approach achieves faster a consensus among end users, stakeholders and healthcare software producers in order to develop interoperable solutions and services. That is the main objective of this paper: to demonstrate how this approach will help decision makers, project implementers as well as stakeholders to define and implement solutions that fit with interoperability requirements in a transparent and formal process. After a presentation of the concepts used, the methodology is presented, and experiences and best practices will complete the demonstration.

BACKGROUND

It is now consensually accepted that the interoperability concept [IEE[1], 1990] covers multiple dimensions (interoperability layers as a part of an interoperability framework) that were well described in the Antilope project (2015) and synthesized in Figure 1.

Figure 1. Refined Interoperability model (Antilope, 2015)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Interoperability is defined in the layers mentioned below:

- The legal and regulatory layer that defines the legal constraints and rules for establishing lawful health information exchange;
- The policy layer that defines the needed organizational rules, as well as, how the stakeholders will work together when the patient provides his (her) consent;
- The care process layer is aligning with the medical and social processes;
- The information layer contains both semantic and syntactic interoperability rules (terminologies, coding, etc.);
- The applications layer describes the import export, messaging and display mechanisms of healthcare information;
- The infrastructure layer describes communication and protocols layer (networks, etc.).

A foreseen interoperable implementation has to consider all these layers in order to manage to have a successful deployment.

The transversal layers, security, standards and profiles are applicable for all the interoperability layers by addressing the complexity of such projects. Indeed, to bridge all these layers harmoniously, it requires a careful selection of security rules, standards and integration profiles to be used in order to avoid disruption or gaps in the chain of trusts or information exchanges. To minimize such drawbacks, the use case driven approach accompanies the designer and architect of the system during all the phases of the project.

The ISO DTR-28380 Part 3 (ISO, 2014) document proposes concepts and definitions, as well as, a process for deploying IHE profiles. This paper espouses and accepts the result of this approach and provides guidelines that will facilitate and support Information Technology (IT), eHealth industry, healthcare professionals and systems architects and engineers to implement and deploy their own project properly.

A use case describes in natural language the needs and requirements from an end-user point of view, clinicians, or patients, and provides quick access to underlying information related to deeper details or technical specifications. However, the description needs to follow a formalized template that provides proper guidance (eStandards, 2016). From the use case, we derive the realization scenarios which describe in more detail and translate into technical processes the steps to be supported by the systems or solutions (eStandards, 2016). Here again, a specific template (eHealth, 2015) guides the architects and facilitates the description of the implementation they want to do.

From one use case, one or more realization scenarios can be derived:

- The use case is described at high level, and sometimes it is better to cut it into one or more realization scenarios for better control of the implementation step as well as for priority related reasons (important steps, timing of procedural steps, mandatory actions, alternative inputs or outputs, etc.);
- Two or more interoperability implementations are technically feasible to support the transition phase, especially in the cases of technological gaps between existing systems and future implementations.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

USE CASE DRIVEN APPROACH

The objective for designers and architects is to design rapidly a robust and implementable infrastructure based on pre-existing best practices. The existing landscape of standards and profiles is vast, and it is not easy to find the proper path in this "forest" when not sufficiently familiar enough or when new standards appear and request recognition.

A use case driven approach facilitates the access to established standards because:

- This approach has been now deployed for more than 15 years with success;
- References, documentation, guidelines are now largely available and accessible (IHE, 2016a; IHE, 2016b; Jacobson, Christophe, Johnsson & Overgaard, 1992; Antilope, 2015);
- A repository of use cases (developed in eStandards project (2016)) is easily accessible providing examples of interoperable realization scenarios and set of profiles and standards;
- A Conformity assessment scheme (IHE, 2015) with a process description and testing tools are available for supporting projects that want to assess their solutions before final deployment.

The approach is supported by the four following methodological steps:

Step 1: Describe the use cases in natural language.
Step 2: Prioritize and select the most higher needs to start.
Step 3: Specify the realisation scenarios and define the eHealth architecture support of the use cases and select standards and profiles based consultation of the use case repository.
Step 4: Select testing tools and analyse when a conformity assessment process can be introduced in the project.

Step 1: Describe the Use Cases

The Refined European Interoperability Framework (ReHIF) developed in Antilope project (Antilope, 2015) as well as the support tools to describe methodologically the use cases were validated by the eHealth Network[2] in November 2015. The ReEIF proposes:

- A set of use cases distributed among ten domains (medication, radiology, laboratory, Patient summary, referral and discharge reporting, participatory healthcare, telemonitoring, multidisciplinary consultation, antenatal care and public health);
- A template that allows harmonization of the use case description that provides a common language and facilitates their reusability among stakeholders (see Figure 2).

Additionally, in order to facilitate adoption and taking-up, a use case data repository (eStandards, 2016) was developed from the eStandards project. A multi-criteria research feature allows the end-users access similar use cases to the ones they intend to implement.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Set of the ReEIF use cases

#	Medical domain	Description	Scale
1	Medication	e-Prescription and e-Dispensing	1a) Cross-border 1b) National/Regional 1c) Intra-organisational 1d) Citizens at home
2	Radiology	Request and results sharing workflow for radiology	2a) National/Regional 2b) Intra-organisational
3	Laboratory	Request and results sharing workflow for laboratory	3a) National/Regional 3b) Intra-organisational
4	Patient Summary	Patient Summary sharing	4a) Cross-border/International 4b) National/regional 4c) Citizens at home
5	Referral- and Discharge reporting	Cross-enterprise Referral and Discharge Reporting	National /Regional 5a) Referral of patient from primary to secondary care 5b) Discharge report from secondary care
6	Participatory healthcare	Involvement by chronic patients in electronic documentation of healthcare information	Citizens at home
7	Telemonitoring	Remote monitoring and care of people at home or on the move using sensor devices	Citizens at home
8	Multidisciplinary consultation	Medical Board Review	National/Regional
9	Public Health	Immunization	National/regional Intra Organizational
10	Antenatal care	Antenatal care	National/Regional

Each use case provides information on the:

- **Relevance:** Describes in natural language the needs;
- **Domain:** One of the ten domains;
- **Scalability:** Use case implemented at the local, regional/national, European or international levels;
- **Context:** Challenges, ehealth strategy or objectives to be reached, benefits, etc.;
- **Information:** Medical information that will be shared among stakeholders;
- **Participants:** Healthcare professionals and patient involved in the use case;
- **Process flow:** Describes the information flow among participants;
- **Associated realization scenario(s):** One or more realization scenarios answer the needs and are described in more detail in the allocated template.

  Example: Relevance of the Patient Summary at the national/regional scale.

"For the exchange of medical information about a patient, a Patient Summary provides the participants in a healthcare pathway with the 38 basic medical background information. In collaborative healthcare, in the transfer of a patient to another hospital, and in a multidisciplinary board review, the Patient Summary functions as the standardized information set of medical information. The structuring of basic information such as current medication, allergies, advance directives, diagnoses, and therapies allows the different healthcare information systems to absorb the information from any other healthcare information system. The treatment of patients without proper medical background information is hazardous and should be avoided. This Use Case proposes a way towards solving this problem."

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Step 2: Prioritize the Use Cases

Best practices showed that the best way to deploy healthcare information exchange is to define an iterative approach based on use case selection. Therefore, the set of use cases will be implemented incrementally and step by step regarding a list of criteria established by each project based on its priorities and goals. For each functional or technical criterion, a weighting is defined related to their relevance and impact on the country or project objectives. For example, the list of criteria could contain:

- Health challenges:
  - **Benefit for the population:** Increase of the quality of life and the patient safety, health effectiveness, decrease of cost of care, etc.;
  - **Benefit for the healthcare professionals:** Better access to the patient medical information, to the peer expertise, to the decision support, etc.
- Maturity of the use case:
  - Number of deployed equivalent projects over the world;
  - Level of complexity;
  - Use case is requested by several channels (end-users, projects, national programs, etc.);
- Technical criteria:
  - Maturity and availability of technologies and standards that will support the use case;
  - Derivation of the use case on realization scenario is available;
  - Estimated timescale and cost for the use case deployment;
  - Infrastructure relationships between use cases that will be implemented;
- Cost Criteria:
  - Cost of implementation versus available funds (Capital expenses – CAPEX);
  - Cost of operation versus generated income or savings (operational expenses - OPEX);
  - Budgeting – funding restrictions.

Of course, this list is not exhaustive. Other criteria can be found according to the project's context of a specific country or region.

Step 3: Specify Realization Scenarios

One or more realization scenarios are closely coupled to a specific use case. They describe high-level interoperability implementation and more specifically the systems' actors and their interaction using standards and profiles. They are formalized in a template which includes:

- The related use case that the realization scenario implements;
- The context that describes the interoperable architecture and specific technical requirements;
- The actors that are systems or solutions involved in the workflow;
- The transactions that correspond to: the messages, documents exchanging or shared data between systems;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- The technical process flow(s) that can be represented by an interaction diagram or in a description depicting actors and transactions in the flow;
- The associated profiles and standards underneath;
- Other possible issues raised during the specification process which may cover security issues, options not yet chosen, architecture issues, custom interfacing with legacy systems and much more.

In the eStandards use case repository, the use cases are associated with the realization scenarios offering the possibility to easily access a specific selection of profiles and standards that serve a scenario. Interaction diagrams or interaction descriptions indicate the way to implement the scenario and can be adapted for specific usages.

The benefits for the implementers are quite obvious:

- Provide guidance to implementers who are not familiar with the ehealth standard complexity;
- Provide examples of architecture implementation proposals;
- Provide quick access to the specifications of profiles and standards;
- Provide the list of European Profiles eligible for referencing in public procurement (European Commission, 2015).

After having chosen the set of profiles and standards that are already available, a gap analysis should be performed in order to check if all the needs of end users are well covered. If it is not the case, the implementers shall select additional messages from standards already selected or add new profiles or standards.

The use case data repository demonstrates that several use cases can benefit from the same set of IHE profiles and therefore supported by the same set of services and infrastructure which optimizes their deployment.

Step 4: Select Testing Tools

Tooling for interoperability is becoming more and more available to the ehealth ecosystems increasing the level of quality of the ehealth project. Two types of tests can be performed:

- The conformance testing checks if the messages or documents provided by the systems conform to given specifications; conformance testing can be done with simulators to simulate processes and actors and validators to validate input and outputs (documents, messages, tokens, etc.);
- The Interoperability testing checks that at least two systems are exchanging correctly information using specific protocols/profiles or profiles/standards in a given workflow. Interoperability testing can be achieved in testing events either face to face or in virtual environments.

Table 1 presents different categories of tools as examples (not exhaustive).

A list of tools is available from the Antilope project deliverable 3.1 (Antilope Testing Tools - WP3.1, 2015). The IHE Gazelle environment proposes a large number of open source tools that allow any implementer to test the conformity of their development using the Gazelle external validation[8] (see Figure 3).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Different categories of tools as examples

Type of Tools	Tools	Examples
Specification	Localise the standards for a project, create analytical business rules specifications in documents	Art Décor for HL7 CDA[3] documents (Boufajha, Poiseau, Tomazon & Bergé, 2015)
Validator or conformance tester	Validation of the conformity to the standard	Gazelle™[4] HL7 validator DICOM[5] validator Assertion validator IHE XD*[6] metadata
Simulator	Play the role of the systems performing the behavior of one or several actors	Gazelle™ simulators for IHE profiles
Management	A tool that manages the testing session including the validation report, the test plan and registers the systems under test and their performed tests for traceability, the testers identification and the monitors identification	Gazelle™ TestBed
Proxy	A tool that captures messages exchanged between systems or actors. The validation of the messages can be performed afterward by using the other tools	Gazelle™ proxy
Test data generator	Demographic data server (DDS) Private Key Infrastructure (PKI)for testing	Gazelle™ DDS
Other Support tools	Libraries, examples of implementation, issue tracker, terminology servers, and much more	

Note:

HL7 Clinical Document Architecture standards specify the structure of a clinical document such as Patient Summary, ePrescription or any other document clinical specialties.

DICOM is the standard for exchanging or sharing medical imaging developed by NEMA[7] (the Association of Electrical Equipment and Medical Imaging Manufacturers)

IHE XD* are IHE integration profiles are used for exchanging or sharing medical documents between healthcare enterprise. The specifications describe actors (supported by systems), transactions (messages that are exchanging among actors) and metadata.

Gazelle is a test bed developed by IHE

Figure 3. Gazelle external validation (IHE International)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This interface offers testing tools for the most deployed standards and provides opportunities for testing implemented specifications of documents or messages for projects such as:

- **epSOS[9] (Europe):** Assertions, audit messages, HL7, XD* metadata[10], CDA documents;
- **CI-SIS[11] (France):** Documents in CDA;
- **eSens[12] (Europe):** eID[13], REM[14], and SMP[15];
- **EFA[16] (Germany):** EFA CDA documents and HL7v2 messages.

IHE has established an international conformity assessment scheme (IHE, 2015) where several of the most popular profiles used in an exchanged or shared infrastructure are selected for thorough testing based on ISO 17025 standard (IS0, 2005): vendors can sell products after having successfully passed a conformance assessment, and then stakeholders can choose to trust high-quality interoperable solutions that abide by those conformity assessment schemes such as the one globally proposed by IHE International.

BEST PRACTICES AND EXPERIENCES

In Europe, several countries have selected use cases related to:

- Patient Summaries and discharge letter for the coordination of the Patient care (pathways) between the hospital and the General Practitioners (GP); This is the case in Luxembourg, France, The Netherlands, Austria, Veneto region; Belgium and other countries and regions;
- eDispensation to have a view on what are the current dispensations for a patient and to analyze the economic impact at the national/regional level (France, Greece, Croatia, etc.);
- ePrescription in Nordic countries (Denmark, Sweden) and Greece, Croatia.

Other use cases are also be considered such as laboratory reports (Luxembourg in concept design progress, Austria, etc) or radiology reports (regional tele Imaging projects in different countries). With the development of the Telemedicine, new challenges are raised with a high demand for Interoperability solutions for:

- Patient at home such as chronic patients (diabetes, COPD, allergies);
- Teleconsultation, tele-expertise or telemonitoring.

The healthcare infrastructures that are emerging shall obviously serve all kinds of use cases or needs. This model can be replicated in any circumstance across the healthcare delivery.

During the past years, European Member States experimented cross-border healthcare exchanges based on XDS architecture, exchanging Patient Summaries, ePrescription and eDispensation in the epSOS project (2015). With the EXPAND project (2015), the demonstration of exchanging medical data was a success, and several lessons were learned (EXPAND project, 2015).

In Table 2, the main issues to be analyzed when setting up a new Health Information Exchange facility are presented.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Main issues when setting up a new health information exchange facility

Layers	National	Cross Border
Legal and regulatory	Verify that the legal framework is relevant for sharing or exchanging medical data (confidentiality and privacy, patient consent, patient and healthcare professionals directories)	Compliant to the Patient' rights Directive (European Parliament, 2011) and European data protection directive. Overview on General Data Protection Regulation (GDPR) (European Commission,2012) A country shall verify that the legal context is compliant with the legal framework developed in epSOS project and updated by the eHN Country shall sign the multilateral agreement (MLA)
Policy and Care process	Agreement between healthcare organizations is in place describing the responsibilities related to the care process and the creation/consumer of the shared documents, etc.	Guideline on Organisational Framework (eHealth Network, 2015)
Information	National/Regional Interoperability framework that includes terminologies, medical documents specifications, identification of objects.	Patient Summary and ePrescription guidelines; Master Value set Catalogue, NCP specifications
Technical services	National/regional infrastructure and services connected to the cross-border infrastructure (interfaces)	National Contact Point (NCP) reference eHealth DSI (eHealth Digital Service Infrastructure
Testing services	National/Regional Conformity Assessment Scheme	European Conformity Assessment Scheme (in progress)

CONCLUSION

The interoperability use case driven approach demonstrates its robustness, and it is now used as a basis when stakeholders want to deploy an interoperable project in eHealth.

It will facilitate the access to the technical specifications that exist in standards and integration profiles. The objective for the next years is to get more harmonization over Europe even if the complete standardization in interoperability is not fully realistic due to specificities in each country or region. It will provide more transparency and openness that will encourage stakeholders and patients to collaborate and to share their experiences. It will also allow implementers and vendors to become more extrovert and create products and services that could apply to more than specific and closed markets.

Use cases and realization scenarios that are supported by the Use Case repository, are defined at high level (or large grain) in order to avoid the multiplicity of use cases that will be not easy to manipulate or not easy to map with operational projects by the stakeholders. It will be efficient to link this approach with the approach described in ISO DTR 28380 (2014) that presents more granular technical use cases directly linked to only one IHE profile.

This approach also provides support to contracting authorities and stakeholders to provide better specifications in their tenders. Next steps will be a better deployment of all these available guidelines and tools at the European level for fostering the harmonization between countries in order to facilitate cross-border exchanges. New use cases will benefit from this approach and the first implementation of existing use cases shall be developed much faster than in the past and will complement and meet the user needs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Antilope project. (2015). Advancing eHealth Interoperability. Retrieved from https://antilope-project.eu

Antilope Testing Tools - WP3. 1 (2015). Retrieved from https://www.antilope-project.eu/wp-content/uploads/2013/05/D3.1-Testing-tools-overview-v1.1.pdf

Boufahja, A., Poiseau, E., Thomazon, G., & Bergé, A. (2015). *Model-Based analysis of HL7 CDA R2, Conformance and requirements coverage*. IHIC.

eHealth Network. (2015). *Guideline on an organizational Framework for eHealth National Contact Point*. Retrieved from http://ec.europa.eu/health/ehealth/docs/ev_20151123_co01_en.pdf

eHealth Network. (2015). *Refined eHealth European Interoperability Framework. eHN meeting documents.* Retrieved from http://ec.europa.eu/health/ehealth/docs/ev_20151123_co03_en.pdf

epSOS project. (2015). *Smart open services for European Patients*. Retrieved from http://www.epsos.eu

eStandards project (2016). eHealth stands and profiles in Action for Europe and Beyond. Use Case repository. Retrieved from http://www.estandards-project.eu

European Commission (2012). *Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).* Retrieved from http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf

European Commission. (2016). *eHealth Network*. Retrieved from http://ec.europa.eu/health/ehealth/policy/network/index_en.htm

EXPAND project. (2015). *Expanding Health Data Interoperability Services*. Retrieved from http://www.expandproject.eu/

IHE International. (2015). *IHE Conformity Assessment Scheme*. Retrieved from http://ihe.net/Conformity-Assessment/

IHE International. (2016a). *Engage HIT Stakeholders in a Proven Process*. Retrieved from http://ihe.net/IHE_Process/

IHE International. (2016b). *External Validation Service Front end*. Retrieved August 30 2016 from https://gazelle.ihe.net/EVSClient/home.seam

ISO TC 215/SC. (2014). ISO DTR 28380-3. Health Informatics - Messages and Communication – IHE Global Standards Adoption Process – Part 3 – Deployment.

ISO. IEC 17025. (2005). General requirements for the competence of testing and calibration laboratories. Retrieved from http://www.iso.org/iso/catalogue_detail.htm?csnumber=39883

Jacobson, I., Christophe, M., Jonsson, P., & Overgaard, G. (1992). Jacobson, I. (1992). Object-oriented software engineering: a use case driven approach. Wokingham, England: ACM Press; Addison-Wesley Pub.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

European Parliament. (2011). Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross border healthcare. *Official Journal of the European Union.* Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:088:0045:0065:EN:PDF

European Parliament. (2015). Commission decision (decision (EU) 2015/1 302 of July 2015 on the identification of "Integrating the Healthcare Enterprise" profiles for referencing in public procurement. *Official Journal of the European Union*. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32015D1302

ENDNOTES

1 Institute of Electrical and Electronics Engineers.
2 The eHealth Network was established by the directive 2011/24/EU on the application of patient's rights in cross-border healthcare gathers voluntary European Member States in order to ensure coordination, coherence and consistency of the work on eHealth Union level and to avoid duplication of work.
3 HL7 CDA: Health Level Seven Clinical Document Architecture
4 https://gazelle.ihe.net/
5 DICOM Digital Imaging and Communications in Medicine
6 IHE XD*Cross Referencing Document star
7 www.nema.org
8 https://gazelle.ihe.net/EVSClient/home.seam
9 epSOS: European Patient Smart Open Services
10 Means cross domain metadata
11 Cadre d'Interopérabilité des Systèmes d'Information de Santé (CI-SIS) http://esante.gouv.fr/services/referentiels/referentiels-d-interoperabilite/cadre-d-interoperabilite-des-systemes-d
12 http://www.esens.eu/
13 eID: electronic identification
14 Registered Electronic Mail (REM)
15 Service Metadata Publishing (SMP)
16 https://de.wikipedia.org/wiki/Elektronische_Fallakte

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 6(3); edited by Anastasius Moumtzoglou; pages 52-62; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 20
Agent-Based Modelling in Multicellular Systems Biology

Sara Montagna
Università di Bologna, Italy

Andrea Omicini
Università di Bologna, Italy

ABSTRACT

This chapter aims at discussing the content of multi-agent based simulation (MABS) applied to computational biology i.e., to modelling and simulating biological systems by means of computational models, methodologies, and frameworks. In particular, the adoption of agent-based modelling (ABM) in the field of multicellular systems biology is explored, focussing on the challenging scenarios of developmental biology. After motivating why agent-based abstractions are critical in representing multicellular systems behaviour, MABS is discussed as the source of the most natural and appropriate mechanism for analysing the self-organising behaviour of systems of cells. As a case study, an application of MABS to the development of Drosophila Melanogaster is finally presented, which exploits the ALCHEMIST platform for agent-based simulation.

INTRODUCTION

The chapter reviews the role of *agent-based modelling* (ABM) in the simulation of biological systems, focussing on the simulation of multicellular systems. Modelling multicellular systems requires tools that can support multi-scale models, where different cells form large-scale, dynamic networked systems – as, e.g., in tissues of cells, organs, and even full embryos – and where both the biochemical reactions that occur inside each cell and the molecules diffusion along the tissue (mediating the interaction among nuclei/cells) can be reproduced (Dada & Mendes, 2011; Deisboeck, Wang, Macklin, & Cristini, 2011).

Despite its recognised value for modelling complex systems (Bonabeau, 2002), and also for modelling biological systems (Holcombe et al., 2012), ABM is still not completely accepted as a suitable modelling

DOI: 10.4018/978-1-7998-1204-3.ch020

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

approach in the literature on multicellular systems—whereas, dually, multicellular systems are not usually considered as a valid application domain in the ABM literature. This provides the first motivation behind this chapter: that is, pointing out the potential of ABM in the context of multicellular systems.

ABM grounds on top of the three fundamental abstractions for multi-agent systems (MAS) (Zambonelli & Omicini, 2004):

- **Agent:** The autonomous source of activity in the system,
- **Society:** A group of interacting agents, and
- **Environment:** The context where agents and societies live and interact. As such, agent-based models are perfectly suited for studying the systemic and emergent properties that characterise a multicellular system, meant to be reproduced *in virtuo*.

ABM looks a promising approach in this context given the direct mapping between agent-based abstractions and multicellular systems: agent-cell (with internal and interacting behaviour), agent society-multicellular system, environment-extracellular matrix.

As an example, *developmental biology* is used as a reference scenario to demonstrate the power of *multi-agent based simulation* (MABS). Developmental biology is a challenging branch of life science that studies the process by which organisms develop, focussing on the genetic control of cell growth, differentiation and movement: as such, it analyses multicellular systems during their formation and development. In particular, the chapter illustrates the application of MABS to the development of *Drosophila Melanogaster* based on the ALCHEMIST agent-based simulation framework (Pianini, Montagna & Viroli, 2013; Montagna, Omicini & Pianini, 2016). In the overall, the contribution of the chapter can be articulated as follows:

1. Reviewing the actual state-of-the-art in MABS applied to multicellular systems biology, as well as making an argument for the value of agent-based approaches in the field;
2. Reporting on a case-study of agent modelling and simulation;
3. Discussing the specific issue of developmental biology – that is, one of the hottest and most challenging field in biology –, and verifying which specific achievements were obtained in the last years by adopting the agent metaphor within biological studies;
4. Providing an example of MABS platform for biological systems; and
5. Supplying biologists with a tool possibly helping them finding insights into the biological phenomena, experimenting working hypotheses, performing the so called *what-if* analysis, and making prediction on system behaviour.

BACKGROUND

The field of research devoted to understanding biological cell functionalities by means of modelling and simulation techniques is called *Systems Biology*. The field originally emerged around year 2000, in particular the pioneering work of Kitano (2002). Systems biology is grounded on the idea that the high complexity of phenomena and mechanisms regulating biological system dynamics cannot be understood by *in-vivo* experiments alone—mostly, due to the limitation of experimental techniques in investigating the coordinated activities among system components. In order to understand biological systems, other

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

concepts and tools are required that could enable the execution of *in-silico* experiments, *i.e.*, modelling and simulation techniques to understand how the integration and interaction of components can yield emergent behaviours.

Systems biology mainly focuses on the analysis of mechanisms at the intracellular scale—such as pathways, genetic, and metabolic regulation. Recently, a massive trend of research moved towards the analysis of the entire multicellular system (Hoehme & Drasdo, 2010), and a new field of research emerged: multicellular systems biology. Accordingly, the remainder of this section provides a description of the main biological facts involving multicellular systems, and points out the critical issues that make such systems particularly interesting from the research viewpoint.

Multicellular Systems Biology

Multicellular systems biology studies multicellular systems. *Multicellular systems are living organisms that are composed of numerous interacting cells*.[1] Humans and most of the animals are multicellular organisms. Research focuses in analysing specific tissues of the entire organism, such as neural system, immune system, adult stem cells etc., that are leading specific biological functions. Moreover, many research initiatives are devoted to analysing systems during embryogenesis and population of cancer cells.

All of the aforementioned systems share some features and properties, which make their understanding particularly challenging. The main property is that they are inherently *multi-scale*—see Figure 1. As such, the global behaviour emerges by mechanisms working across multiple space and time scales. Each scale integrates information from strata above and below, according to the principle of *upward and downward causation* (Uhrmacher, Degenring, & Zeigler, 2005). The building block for the vast majority of mechanisms at each level is the *interactions* among components.

It is possible to identify three main hierarchical levels for multicellular systems (Setty, 2012): molecular, cellular, and tissue.

Intracellular regulatory network controls molecular mechanisms, such as gene expression, receptor activity and protein degradation. The dynamic at intracellular level brings each cell into a state that affects the architecture and function on the cell level. However, the intracellular regulatory processes is not enough for understanding the emergent processes: an individual cell decides on its next developmental step, proliferation, fate determination, and motility, according to intracellular state as well as to constrains deriving from the level above. Cell population acts in concert to develop its anatomy and function.

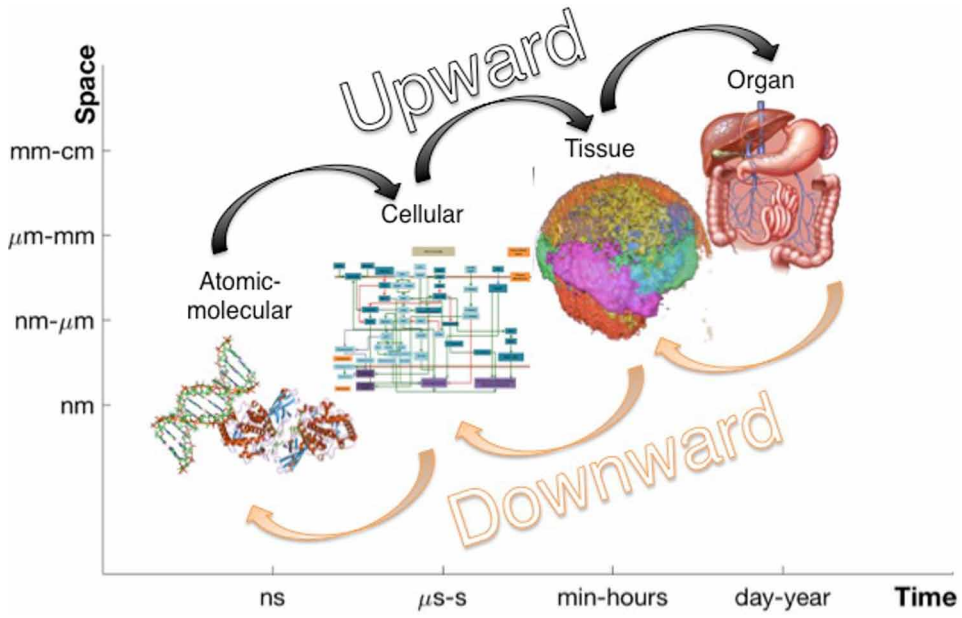
Computational Model Requirements

In the literature, a *middle-out modelling approach* is suggested as a possible technique to reproduce the behaviour of these systems. It starts with an intermediate scale (the cell, the basic unit of life), and it gradually expands to include both smaller and larger scales for reproducing the intra- and inter-scale interactions and integration. It should then be multi-scale in nature, so as to capture several spatial and temporal scales.

Also, it should feature abstractions for modelling diffusion of molecules across cells and in extracellular matrix, for studying the effects of short and long range signals, and for modelling the compartment membrane. Moreover, it should be stochastic, in order to capture the aleatory behaviour characteristic of those systems involving few entities; also, the cell topology should be dynamic, for modelling cell

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A biological system is articulated over at least four levels with different spatial and temporal scales. Each level influences the dynamic at the level above (upward causation) and below (downward causation)



division and movement. Finally, it should be able to capture cell heterogeneity, so as to model individual structures and behaviours of different entities of the biological system.

From the validation perspective, the field is particularly critical since it requires multiple data: molecular data such as gene expression profiles and image data such as spatial-temporal growth pattern.

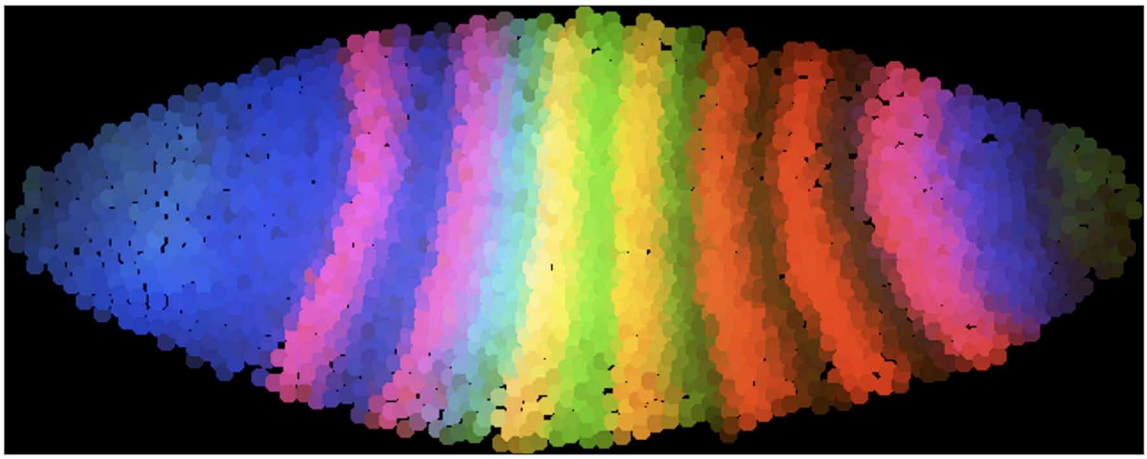
Developmental Biology

Among the different and challenging branches of multicellular systems biology, the chapter focuses on developmental biology, which studies the processes by which multicellular organisms develop. Multicellular organisms pass through similar stages during development. The life of animals begins with the fertilisation of the egg; *fertilisation* stimulates the egg to begin development. A series of extremely rapid and synchronous *mitotic divisions* of cells or cellular nuclei characterises the cleavage stage of development. After that, cells start the process of gene transcription and protein synthesis: cells *differentiate* in that each type expresses a diverse set of genes. Cell diversity is organised into a precise spatial distribution (*embryo regionalisation*), as shown in Figure 2 for the *Drosophila Melanogaster* embryo.

Developmental biology focuses on the genetic control of cell growth, differentiation, and movement, for understanding the mechanisms that make the process of embryo regionalisation so robust, making it possible for an organism to evolve from one cell to the same morphology every time (Hohm & Zitzler, 2009). Answering the big questions of development – such as "How do local interactions among cells and inside cells give rise to the emergent self-organised patterns that are observable at the system level?" – has a double meaning. On the one hand, it should make the fascinating process of multicellular organisms morphogenesis clear. On the other hand, since the molecular and biochemical processes regulating

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The pair-rule gene even-skipped (red) together with hb (blue) and Kr (green) in Drosophila embryo at the cleavage cycle 14A temporal class 8. Reconstructed image from (Pisarev et al., 2009). Embryo name: ba3



cell proliferation, differentiation, survival, and death also play a central role in tissue homeostasis and cancer cells, it could reveal the key mechanisms that drive cancer progression, paving the way towards the identification of new therapies that can control tumour growth.

ABM IN MULTICELLULAR SYSTEMS BIOLOGY

The interdependent nature of multicellular processes often makes it difficult to apply standard mathematical techniques to separate out the scales, uncouple the physical processes or average over contributions from discrete components. (Cooper & Osborne, 2013)

Individual-based models (IBM) are often presented in literature as one of the main alternative to mathematical models based on differential equations. Both cellular automata (CA) and agent-based modelling (ABM) are representative of the class of IBM. Altogether, their computational models are similar, since they are organised around an abstraction of *individual* entities – *cells* in CA, agents in ABM – and the concept of *space*, defining the notion of neighbourhood for each entity. On the other hand, cells in CA are simpler then agents in ABM, since their behaviour is defined with update rules that change their states, switching between boolean values in the basic case, based on the state of the cell itself and of its neighbours. Also, the grid of cells is usually static, so that movement or cell replication is allowed. However, CA cannot model with sufficient accuracy cell proliferation and movement, which are crucial mechanisms especially in tumourogenesis.

This is where the theory of agents and multi-agent systems (MAS) steps in, promoting IBM where individual entities can express a more articulated autonomous behaviour, living, moving, and replicating in a possibly dynamic and complex environment. In the literature, agent-based systems, and MAS in particular, are considered as an effective paradigm for modelling, understanding, and engineering *complex systems* (Omicini & Zambonelli, 2004), and biological systems in particular (Cannata, Corradini, Merelli,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Omicini, & Ricci, 2005), since they provide a basic set of high-level abstractions that make it possible to directly capture and represent the main aspects of complex systems, such as interaction, multiplicity and decentralisation of control, openness, and dynamism (Uhrmacher & Weyns, 2009; Merelli et al., 2007; Macal & North, 2010; An, Mi, Dutta-Moscato & Vodovotz, 2009; Gelfand, 2013).

As from the pioneer work of Bonabeau (2002), an ABM describes the system *from the perspective of its constituent units*. Moreover he states that (Bonabeau, 2002:

The benefits of ABM over other modeling techniques can be captured in three statements:
1. *ABM captures emergent phenomena*;
2. *ABM provides a natural description of a system*; and
3. *ABM is flexible*.

Emergent phenomena result from the interactions of individual entities. By definition, they cannot be reduced to the systems parts: the whole is more than the sum of its parts because of the interactions between the parts. An emergent phenomenon can have properties that are decoupled from the properties of the part. [...] ABM is, by its very nature, the canonical approach to modeling emergent phenomena: in ABM, one models and simulates the behavior of the systems constituent units (the agents) and their interactions, capturing emergence from the bottom up when the simulation is run.

Given that the above definition shares its crucial concepts with the one of systems biology, ABM looks particularly adequate for modelling multicellular systems.

By adopting ABM, biological systems can be modelled as a set of interacting autonomous components, i.e., a set of agents, and their chemical environment can be modelled by suitable environment abstractions, enabling and mediating agent interactions (Weyns, Omicini, & Odell, 2007). In particular, ABM provides a straightforward way to model:

1. The diverse individual structures and behaviours of different entities of the biological system as different agents (heterogeneity);
2. The heterogeneous – in space and time – environment structure and its dynamics; and
3. The local interactions between biological entities/agents (locality) and their environment.

Executing an agent-based model makes it possible to study its evolution through time, in particular:

1. Observing individual and environment evolution;
2. Observing global system properties as emergent properties from agent-environment and inter-agent local interaction; and
3. Performing in-silico experiments.

In ABM each cell is typically modelled as an agent, accounting for independent cell behaviour as well as for cell interactions. In the context of biological system, agent-based models can therefore account for individual cell biochemical mechanisms – gene regulatory network, protein synthesis, secretion and absorption, mitosis and so on – as well as the extracellular matrix dynamic – diffusion of morphogens, degradation and so on – and their dynamic influences on cell behaviour. On top of the cells/agents system, then, there is the whole population, whose behaviour can be observed through simulation, thus show-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ing the relationship between microscopic behaviours of the single agents and macroscopic behaviours of cell population. A further benefit of ABM in multicellular systems biology is that each agent can be associated to a different autonomous behaviour, thus capturing cell heterogeneity that plays a crucial role in a lot of multicellular phenomena.

Different ABM techniques can be found in the literature, such as lattice-based, lattice-free, cellular Potts, lattice-gas, and subcellular element modelling methods (Anderson, Chaplain, & Rejniak, 2007). Typically, they mostly differ for how they model the agent distribution over the environment: for instance, lattice-based models, such as cellular Potts models, organise agents into a lattice structure.

In the Literature

The idea presented in this chapter is supported by related literature that reports about platforms specifically developed for building agent-based models of multicellular systems and running simulation. Generally, a significant number of applications of ABM to biological systems can be found in the literature—even though some of them mainly specifically refer to intracellular networks only (Montagna, Riccim & Omicini, 2008; González et al., 2003; González Pérez, Omicini, & Sbaraglia, 2013; Pogson, Smallwood, Qwarnstrom, & Holcombe, 2006). However, for their very nature, ABM fits multicellular systems, becoming a common method in multicellular system biology (Gorochowski et al., 2012). This is why the remainder of this section will only mention the works related to multicellular systems. In particular, most of the works couple ABM with reaction-diffusion systems that model intracellular or microenvironmental dynamics of substrates.

Platforms

In the following the main ABM platforms are mentioned and shortly described. Since providing details on how to implement a model within each platform is out of the scope of this chapter, interested readers can refer to the original platform documentation.

- **COMPUCELL3D:** (Cickovski et al., 2005) is probably the best example of a hybrid platform that combines a lattice-based ABM (to model cell interactions) and continuous models based on reaction-diffusion equation (to model chemical diffusion). Each compartment – which can be a cell, or even a subcellular compartment – is associated to a set of attributes (like cell type, volume, area), is characterised by a state, and has a behaviour defined by a set of formula changing the value of a variable called *effective energy*, which implements most cell properties, and also governs internal and interacting dynamics. For instance, it accounts for cell adhesion, or elastic constraints that consider volume and surface constraints. In the overall, COMPUCELL3D looks like a very promising framework, whose main limitation is the lack of a realistic model for cell internal behaviour.
- **Morpheus:** (Starruß, de Back, Brusch & Deutsch, 2014) is also is a hybrid framework embodying both ABM and mathematical models such as ordinary differential equation and reaction diffusion systems. Moreover, an integration of the two modelling approaches is possible: each compartment is an agent that, according to the cellular Potts model, defines compartment shape and motility, and whose internal behaviour is usually linked to models of intracellular biological networks specified in the form of ordinary or stochastic differential equations. Morpheus is enhanced with

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

a full-featured graphical user interface that supports the user during model definition and editing as well as during simulation, providing tools for interacting with the simulation and visualising simulation output and results in various ways.

- **ALCHEMIST:** (Pianini, Montagna & Viroli, 2013; Montagna, Omicini & Pianini 2016) is a novel platform based on the concept of nodes/agents, whose behaviour is specified by chemical-like rules. It is particularly suited for modelling biological systems, being its main abstractions very close to the main components of such systems. Details on that will be given in the next section.
- **Other Platforms:**. Beyond the aforementioned platforms, which are the best examples of ABM platforms specifically developed for multicellular systems, other examples are worth to be mentioned shortly in the following. CellSys (Hoehme & Drasdo, 2010) implements a class of lattice-free agent-based models permitting realistic simulations of tissue growth. Currently, the platform seems to be not maintained anymore.

Other platforms are built for developing ABM-MABS in diverse domains, thus without specific abstractions devoted to modelling biochemical mechanisms. For instance, FLAME (Holcombe et al., 2012) implements a pure agent-based model. The behaviour is here modelled in terms of state machines, i.e., a set of states with transition rules or functions that determine how to make the agent evolve. In biology, it has been applied to the modelling of stem cells. Repast (North et al., 2013), MASON (Luke, Cioffi-Revilla, Panait, Sullivan & Balan, 2005), NetLogo (Wilensky, 1999) are the most popular MABS platforms. As such, they provide tools for implementing the typical ABM abstractions. Agent behaviours are defined in different ways: the most general-purpose one is by programming languages (Java for MASON and Repast; a Logo-inspired language for NetLogo) by which any possible behaviour can be given to the agent and to the environment that contains the agents.

Successful Applications

Mostly because it makes it possible to model cell diversity, ABM has been frequently adopted in order to simulate cancer growth, also accounting for the stochasticity of genetic mutations. Wang et al. (2015) provide a review of the most recent ABM approaches developed for studying different aspects of cancer growth. Accordingly, in the following the most representative works showing how ABM can capture the main mechanisms of cancer development are shortly introduced and discussed.

Macklin et al. (2012) describe an agent-based cell model of ductal carcinoma in situ (DCIS), which is a type of breast cancer. Each cell is modelled as a mobile agent that moves according to adhesive, repulsive, and motile forces. Heterogeneity is accounted for by associating each agent to a phenotypic state, and by means of rules that govern phenotypic transitions depending on the internal state of the cell and the local microenvironment. Cellular volume is also reproduced accounting for variations during proliferation and necrosis processes. Such a model is coupled with a reaction-diffusion model for modelling substrate dynamics, namely for describing how cells affect the concentration of molecules in their microenvironment and the dynamic of intracellular molecular signalling. The model predicts that DCIS tumours grow, in agreement with mammographic data.

Olsen and Siegelmann (2013) design an ABM of various cancer types, focussing in particular on cancer growth. It has three levels of representations: the cellular, tissue, and molecular ones.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cells are agents, and a notion of environment is explicitly represented: it mainly models blood vessels that bring nutrients to cells, thus modelling the molecular level. Agents are organised into a three-dimensional grid: the tissue level emerges from the agents and the environment dynamics and interactions. Their behaviour is represented via flow-charts.

Finally, it is worth mentioning a number of other multicellular systems whose mechanics has been investigated in terms of ABM. Gorochowski et al. (2012) analyse the relationship between single-bacterial-cell dynamic and the population level characteristics for studying the collective behaviours observable in many bacterial populations. In Rouillard and Holmes (2014) a case of ABM application to myocardial cells has been reported. It studies the long-term consequences of therapies that alter infart mechanics. These therapies aim at reducing complications of heart attack by altering heart structure, mechanics, and function over time. Walker, Wood, Southgate, Holcombe, and Smallwood (2006) present an extension of the *Epitheliome*, an agent-based model of epithelial cells. There, agents represent individual cells that, in the original *Epitheliome* model, iteratively change their state according to a set of rules based purely on observations and responsible for injecting behaviours into the agents, such as proliferation, intercellular adhesion, migration, and apoptosis. In the last version presented by Walker et al. (2006), a mathematical model is integrated within each agent, to directly model intracellular pathways that cause cell proliferation, thus making a connection between molecular-level events and growth on a tissue level.

Given this big picture of related works supporting the crucial role that ABM has in multicellular systems biology, hereafter we provide details on the computational model of a specific platform and we then show how it applies to modelling a case of embryo development. The goal is to provide readers with an example of what we described until now, an example that can possibly guide interested readers in this fascinating field. Among the tools described before, we choose ALCHEMIST because it is particularly suited to model the case study, and also because, being developed by our research group[2], we are particularly competent on the platform. The example is extracted from a literature paper (Montagna, Pianini, & Viroli, 2012).

AN ILLUSTRATIVE EXAMPLE: THE ALCHEMIST PLATFORM FOR SIMULATING *DROSOPHILA MELANOGASTER* DEVELOPMENT

The ALCHEMIST platform[3] was specifically developed for running models composed by interacting compartments whose internal behaviour is defined by a set of chemical-like rules (Pianini, Montagna, & Viroli, 2013). The basic idea was to overcome the limits we found in other platforms, where the computational model is hard to understand. There, the representation of biological phenomena is far from the real mechanisms actually occurring in biological systems, where everything is a reaction.

The engine is based on two optimised versions (Gibson & Bruck, 2000; Slepoy, Thompso, & Plimpton, 2008) of the popular and successful Gillespie's stochastic simulation algorithm (SSA), originally defined for the simulation of chemical systems (Gillespie, 1977). The computational model was then extended to shift the paradigm from the pure chemistry towards the word of ABM, by specifying the agent (internal and interacting) and environment behaviours through the concept of reactions, properly extended from the original concept of pure chemistry found in Gillespie.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Thus, while ALCHEMIST makes it possible to simulate multiple, separate, interacting, and possibly mobile entities, its approach to ABM is still structured around the concepts of *reaction* and *compartment*. As a result, ALCHEMIST computational model, based on these two key concepts, is ideal for modelling multicellular systems, where the basic mechanisms are reactions, and the basic entities are compartments.

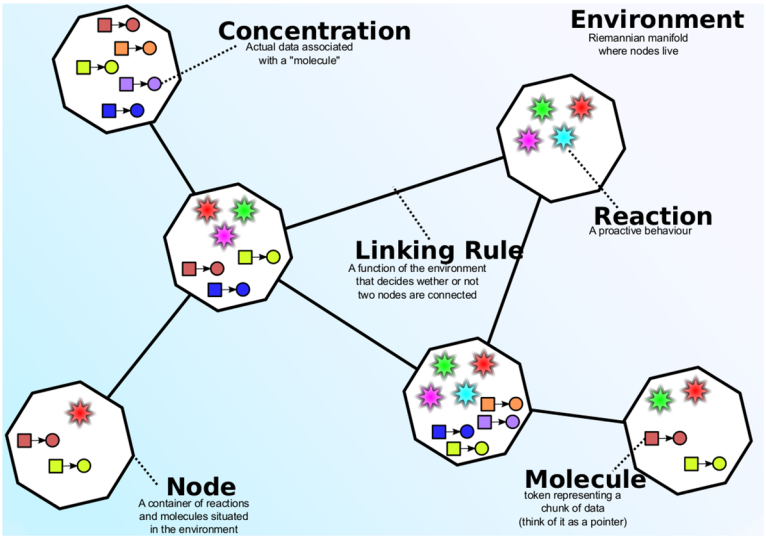
In the following the notions on top of which the ALCHEMIST computational model is grounded are described. Then we show how it is possible to develop a model of a multicellular system using the abstraction provided by ALCHEMIST by using a literature example, the *Drosophila Melanogaster* development shown in a previous work of ours (Montagna, Pianini, & Viroli 2012).

The ALCHEMIST Computational Model

In ALCHEMIST a compartment can represent an agent, and all the possible events in the model should, in the end, break down to a set of reactions (see Figure 3). As such the notion of agent we can deal with, modelling a system with ALCHEMIST is a soft one, *i.e.*, meaning that we can hardly give to agents cognitive capabilities −as it can be required in other simulation domains such as social sciences. However it is perfectly suited for cellular systems. Accordingly, in the following discussion the terms compartment and agent will be used interchangeably as synonyms.

To equip ALCHEMIST with all the abstractions proper of ABM, as first step, the notion of environment, missing in chemistry-derived SSAs, is introduced as a first-class abstraction (Weyns, Omicini & Odell, 2007). The environment has the responsibility to provide, for each compartment, a set of compartments that are its neighbours. The function of the current environment state that determines whether or not two compartments are neighbours can be arbitrarily complicated. Also, it is responsible of exposing possible physical boundaries, namely, to limit movements of compartments situated within the environment.

Figure 3. In the ALCHEMIST computational model, a system is composed of interacting agents whose state is defined by the molecules it contains. The agent state changes over time for the execution of a set of reactions that model the agent behaviour



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The fact that reactions are the only abstraction the modeller can rely upon in order to specify the agent behaviour and to let the simulated system progress does not actually hinder expressivity. In fact, the ALCHEMIST model does not use a strictly-chemical agent behavioural model (White & Pagurek, 1998), rather it generalises the concept of reaction as: *"a set of conditions that, when matched, trigger a set of actions on the environment"*. Accordingly, a condition is a function that associates a numeric value ranging from zero to positive infinity to each possible state of the environment and of agent internal state. Conditions can be everything: for instance, in computational biology, with conditions the modeller can verify the number of neighbouring compartments, the presence of specific molecules inside the agent or in the surrounding environment. If the value associated to condition is zero, the event cannot be scheduled; otherwise, it is up to the reaction to interpret the number. In this framework, actions are arbitrary changes to the environment and/or to agent state. Examples can be the movement of agents, the creation or destruction of molecules, the creation of new agents.

Reactions compute their own expected execution time. Such a putative time is generated by a function ("rate equation" in the following) taking as input all the values generated by conditions and a time distribution. The engine may require putative time to be updated in case that:

1. The reaction has been executed, or
2. Another reaction that the reaction depends on has been executed.

The injection of a time distribution as a parameter for the rate equation makes it possible to model a whole set of events that are not exponentially distributed: imagine, for instance, a burette dropping some quantity of reactant in a compartment every fixed time interval. Such an event is clearly not memoryless, and can be modelled in the proposed framework with a reaction having no condition and a single action increasing the reactant concentration in the compartment. Such a reaction can be fed with a Dirac comb, and the resulting system would seamlessly mix exponential and non-exponential events.

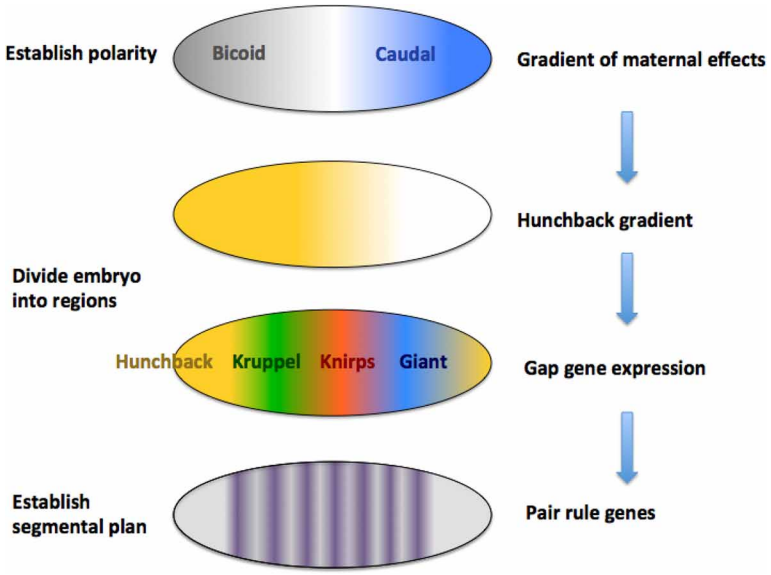
A Case Study in Developmental Biology

To show how to build a model of a multicellular system on top of ALCHEMIST, hereafter we present a case of embryo development. We inspire our discussion to a model of *Drosophila Melanogaster* development presented in Montagna et al. (2012). Drosophila is one of the best-known multicellular organisms. The model reproduces the process of *Drosophila* development from the fertilised egg until the pattern formation of the products of a set of crucial genes called gap genes—*hb*, *kni*, *gt*, and *Kr*.

A schematic representation of what happens in the embryo in the first hours after fertilisation is shown in Figure 4. The schema sketches (from top to bottom) how the concentration of diverse gene products (mRNA and proteins) varies: starting with a unique cell with a polarised distribution of maternal proteins, the downstream transcription of different genes divides the embryo in regions that are stripe-like patterns of expression that strongly characterise the *Drosophila* embryo—as shown in Figure 2. The formation of this pattern is crucial since these segments are closely related to the final anatomical structure. At the same time, fast compartment divisions happen so that the embryo – at that stage – consists of about 6000 separate cells. Understanding how it is possible to create such a precise spatial distribution of gene products, thus generating an exact arrangement of different cell types, is the fascinating challenge of developmental biology, for which the help of modelling techniques is crucial.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Hierarchy of genes establishing the anterior-posterior body plan of the Drosophila Melanogaster embryo



The Model

The whole embryo is modelled as a unique elliptic environment. Each cell is a compartment/agent situated inside the environment. It carries different types of molecules, can move and divide. The 2D continuous environment regulates the position of agents, specifically during movement and growth that must occur within its boundaries. Moreover it is filled with diffusing proteins. The mechanisms explicitly modelled are:

1. Molecule interactions,
2. Molecule diffusion,
3. Agent divisions, and
4. Agent migration.

Molecule Interactions

According to the literature, the gene regulatory network involving gap genes and maternal protein is modelled as a set of reactions where each gene is activated or inhibited by the others. In particular the four gap genes are the molecular actors of this model. They are modelled in two different states, active and inactive (for instance, for the *hb* gene, we have respectively gHb0 and gHb1). Their products are also explicitly modelled. For instance, for the *hb* gene, we have pHb. To exemplify this rule we provide here an example through the reaction that models the inhibition of the *hb* gene by the *Kni* protein: the conditions is that the agent comes with at least one *Kni* protein and the *hb* gene active; the action changes the state *hb* gene from active to inactive. The reaction looks like:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

[pKni + gHb1] --> [pKni + gHb0]

The whole network, including all the interactions among gap genes, and their activation by maternal proteins is fully described by Montagna et al. (2012).

Molecule Diffusion

Diffusion occurs from / to agents and inside the environment along the x and y axis. Molecules move from one agent into the environment, or, the other way round, from one location of the environment into a neighbouring agent. Depending on the type of diffusion (passive or active transport), conditions – for this reaction to be scheduled – are diverse, such as the presence of membrane proteins or of energy in the form of ATP. A simple example follows:

[pHb in cell] --> [pHb in environment]

Agent Division

It is modelled as a chemical-like reaction, whose condition is given by the maximum number of other agents in the neighbourhood and whose action is the creation of a new agent. The new agent is situated close to the dividing one in a casual direction, provided that an other agent does not occupy the new position, and owns half of its molecular content. The rate of compartment division is determined according to the rate observed in the real system. In particular we refer to data reported in Montagna et al. (2010). Since experimental data show a given synchronisation between dividing cells, the phenomena was modelled through a non-Markovian reaction. The reaction looks like:

NumberOfNeighbors < 6 --> Division

Agent Migration

Movement of agents is based on biomechanical forces of repulsion among neighbouring agents: if two agents are closer than a distance – given as a parameter – that models agent surface (spatial extension), a repulsive force is generated and a new position for them is computed. The nearer two agents are, the stronger is the repulsion. Moreover they are forced to remain within the membrane-delimited area so as to filling pretty homogeneously the available space.

Simulation Results

The initial set up for simulations resembles the initial state of the embryo, where only one compartment is present, and where maternal proteins form a gradient in the environment as from experimental data (Pisarev et al., 2009). Simulations results are shown in Figure 5-6-7. They are evaluated observing the time evolution of the compartment number and of the gene expression pattern—see Figure 8. Figure 5 shows the initial condition with only one agent and the egg polarised by maternal effects localised in the extreme pole: *bcd* on the left and *cad* on the right. During the first minutes of simulation only agent divisions occur so as to fill the whole maternal cell at the end of cleavage cycle 9, *i.e.*, after 9 compart-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ment divisions, with around 250 agents (half of the total 500), as shown in Figure 6. Finally, Figure 7 reproduces the expression pattern of the four gap genes, whose spatial organisation is compared with experimental data of Figure 8. Cells are coloured of yellow, red, blue, and green—if, in order, they express *hb*, *kni*, *gt* and *Kr*, and their size is proportional to the protein concentration.

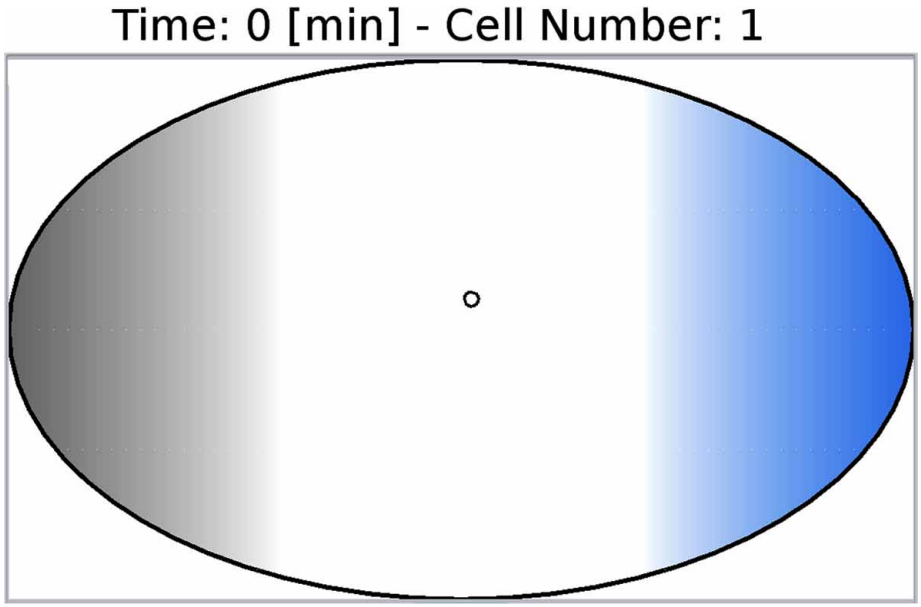
FUTURE RESEARCH DIRECTIONS

Directions for future research can be summarised into two main issues. The first theoretical issue is to refine the way multicellular systems are mapped into an agent-based model and the way simulations are executed, analysing the algorithms adopted to make time proceeds and events execute accordingly.

Secondly – to demonstrate the feasibility and effectiveness of ABM and MABS – our next focus is to apply the ALCHEMIST framework to the analysis of the main morphogenetic signalling pathways and of their interactions. This is a crucial problem in developmental biology, given that they are also involved in tumour growth. Since the integrated phenomenon is highly complex, involving a large set of mechanisms mediating cell-to-cell and intracellular interactions, literature evidences (Dada & Mendes, 2011; Deisboeck et al., 2011) show that novel algorithms and tools from computer science can significantly contribute in supporting experimental biologists to ravelling out such complexity.

To make the model description more clear, and to support the reproducibility and evaluation of the model, we plan to adopt protocols and standards, such as the ODD (Overview, Design concepts, and Details) one (Grimm et al., 2010), which should make agent-based models easier to write and read—and, more generally, the ABM approach easier to embrace for the multicellular systems biology community.

Figure 5. The initial configuration of the simulation is characterised the fertilised egg, characterised by only one compartment and by the gradient of maternal product, bcd (grey) and cad (light blue), in the environment



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. At time 72, after 9 nuclear divisions, the number of compartments is around 2^9. The image shows half of them, being the environment 2D.

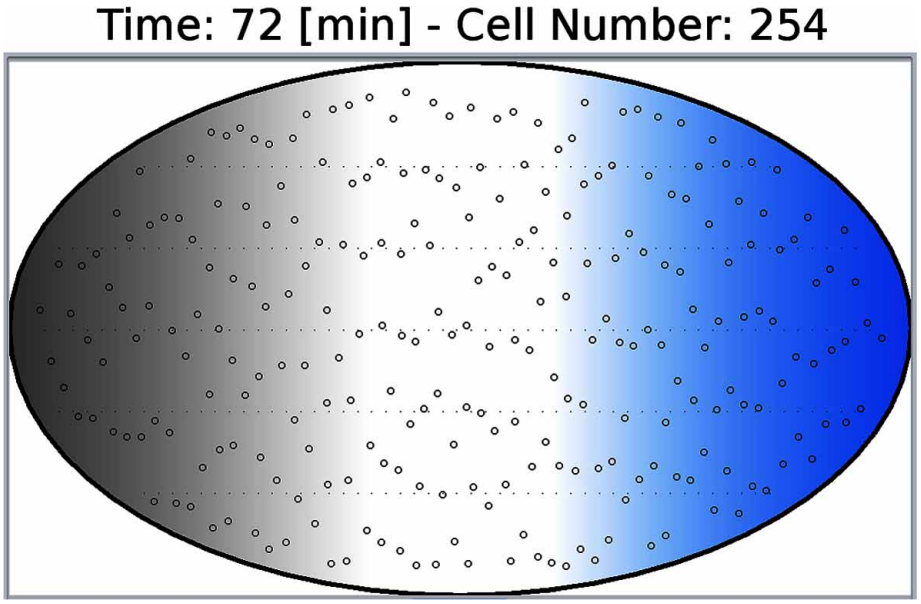
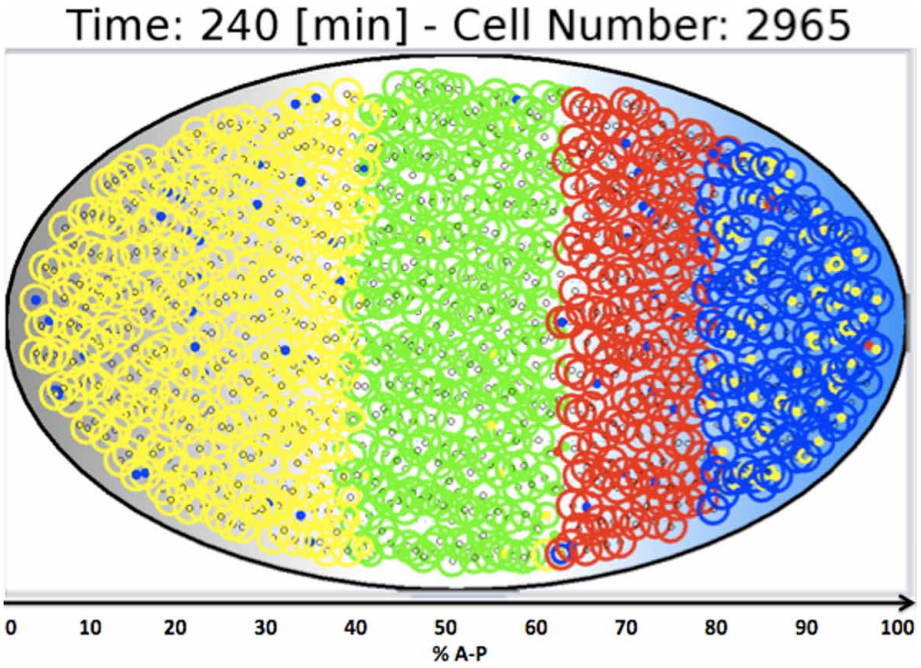
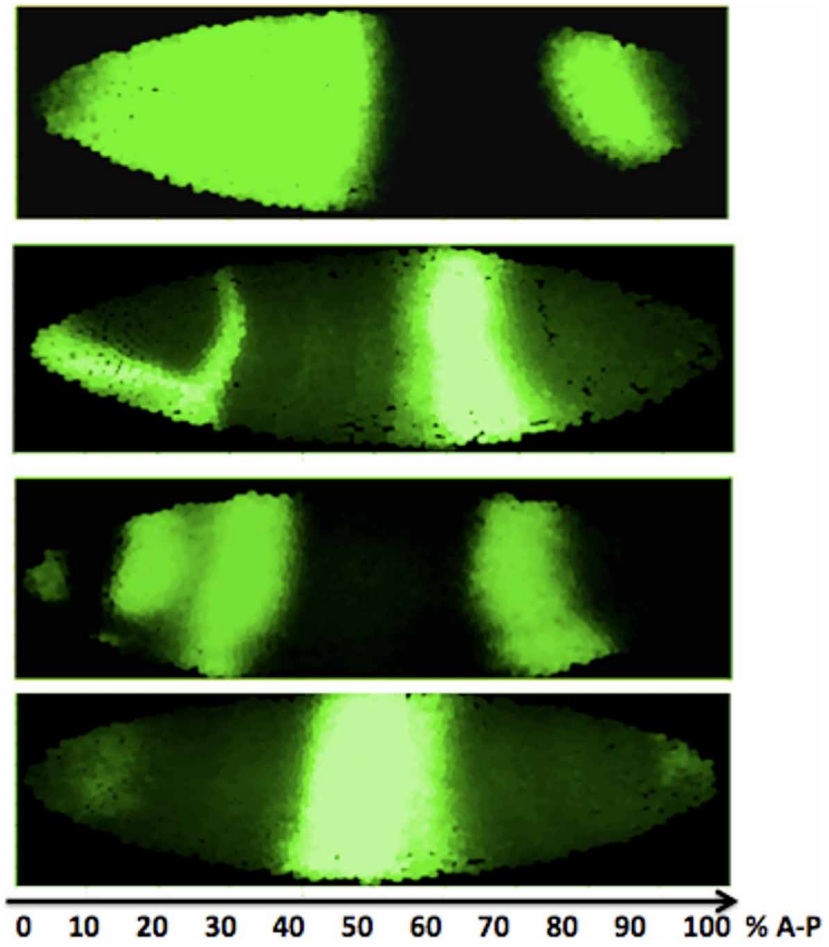


Figure 7. Simulation results for the four gap genes hb (yellow), kni (red), gt (blue), Kr (green) at a simulation time equivalent to the eighth time step of cleavage cycle 14A



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. The experimental data for the expression of (from the top) hb, kni, gt, Kr at the eighth time step of cleavage cycle 14A
Source: Surkova et al., 2008 ©Maria Samsonova and John Reinitz



CONCLUSION

The goal of this chapter is to present a promising and challenging application domain for ABM: multicellular systems biology. The constituting characteristics of multicellular systems in fact highlight a clear mapping between the real system entities and the ABM abstractions. This makes building models and interpreting simulation results intuitive. For this reason we think that ABM is particularly suited for multicellular systems biology. To support our vision, we recollect the literature reporting on successful applications of ABM in this field, as well as a list of ABM platforms specifically developed for building models of interacting cells. To help the reader to better understand better the big picture, we finally choose one platform, ALCHEMIST, and motivate its adoption for ABM of multicellular systems. Finally we describe an example from the literature, which specifically shows how to model a system of interacting cells.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

An, G., Mi, Q., Dutta-Moscato, J., & Vodovotz, Y. (2009). Agent-based models in translational systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *1*(2), 159–171. doi:10.1002/wsbm.45 PMID:20835989

Anderson, A. R. A., Chaplain, M. A. J., & Rejniak, K. A. (Eds.). (2007). *Single-cell-based models in biology and medicine*. Birkähuser Basel; doi:10.1007/978-3-7643-8123-3

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proc. of the National Academy of Sciences, 99*(3), 7280-7287. doi:10.1073/pnas.082080899

Cannata, N., Corradini, F., Merelli, E., Omicini, A., & Ricci, A. (2005). An agent-oriented conceptual framework for Systems Biology. In E. Merelli, P. P. González Pérez & A. Omicini (Eds.), Transactions on Computational Systems Biology III (Vol. 3737, pp. 105–122). Springer. 8 doi:10.1007/11599128_8

Cickovski, T. M., Huang, C., Chaturvedi, R., Glimm, T., Hentschel, H. G. E., Alber, M. S., Glazier, J. A., Newman, S. A. & Izaguirre, J. A. (2005). A framework for three-dimensional simulation of morphogenesis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2*(4), 273–288. doi:10.1109/TCBB.2005.46

Cooper, J., & Osborne, J. (2013). Connecting models to data in multiscale multicellular tissue simulations. *Procedia Computer Science*, *18*, 712–721. doi:10.1016/j.procs.2013.05.235

Dada, J. O., & Mendes, P. (2011). Multi-scale modelling and simulation in systems biology. *Integrative Biology*, *3*(2), 86–96. doi:10.1039/c0ib00075b PMID:21212881

Deisboeck, T. S., Wang, Z., Macklin, P., & Cristini, V. (2011). Multiscale cancer modeling. *Annual Review of Biomedical Engineering*, *13*(1), 127–155. doi:10.1146/annurev-bioeng-071910-124729 PMID:21529163

Gelfand, A. (2013). The biology of interacting things: The intuitive power of agent-based models. *Biomedical Computation Review*, *1*, 20–27.

Gibson, M. A., & Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *The Journal of Physical Chemistry A*, *104*(9), 1876–1889. doi:10.1021/jp993732q

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, *81*(25), 2340–2361. doi:10.1021/j100540a008

González Pérez, P., Cárdenas, M., Camacho, D., Franyuti, A., Rosas, O., & Lagúnez-Otero, J. (2003). Cellulat: An agent-based intracellular signalling model. *Bio Systems*, *68*(2–3), 171–185. doi:10.1016/S0303-2647(02)00094-1 PMID:12595116

González Pérez, P. P., Omicini, A., & Sbaraglia, M. (2013). A biochemically-inspired coordination-based model for simulating intracellular signalling pathways. *Journal of Simulation*, *7*(3), 216–226. doi:10.1057/jos.2012.28

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gorochowski, T. E., Matyjaszkiewicz, A., Todd, T., Oak, N., Kowalska, K., Reid, S., ... di Bernardo, M. (2012). 08). BSim: An agent-based tool for modeling bacterial populations in systems and synthetic biology. *PLoS ONE*, *7*(8), 1–9. doi:10.1371/journal.pone.0042790 PMID:22936991

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. doi:10.1016/j.ecolmodel.2010.08.019

Hoehme, S., & Drasdo, D. (2010). A cell-based simulation software for multi-cellular systems. *Bioinformatics (Oxford, England)*, *26*(20), 2641–2642. doi:10.1093/bioinformatics/btq437 PMID:20709692

Hohm, T., & Zitzler, E. (2009). Multicellular pattern formation. *IEEE Engineering in Medicine and Biology Magazine*, *28*(4), 52–57. doi:10.1109/MEMB.2009.932905 PMID:19622425

Holcombe, M., Adra, S., Bicak, M., Chin, S., Coakley, S., Graham, A. I., ... Worth, D. (2012). Modelling complex biological systems using an agent-based approach. *Integrative Biology*, *4*(1), 53–64. doi:10.1039/C1IB00042J PMID:22052476

Kitano, H. (2002). Systems Biology: A brief overview. *Science*, *295*(5560), 1662–1664. doi:10.1126cience.1069492 PMID:11872829

Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., & Balan, G. (2005). MASON: A multiagent simulation environment. *Simulation*, *81*(7), 517–527. doi:10.1177/0037549705058073

Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation*, *4*(3), 151–162. doi:10.1057/jos.2010.3

Macklin, P., Edgerton, M. E., Thompson, A. M., & Cristini, V. (2012). Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression. *Journal of Theoretical Biology*, *301*, 122–140. doi:10.1016/j.jtbi.2012.02.002 PMID:22342935

Merelli, E., Armano, G., Cannata, N., Corradini, F., dInverno, M., Doms, A., ... Luck, M. (2007). Agents in bioinformatics, computational and systems biology. *Briefings in Bioinformatics*, *8*(1), 45–59. doi:10.1093/bib/bbl014 PMID:16772270

Montagna, S., Donati, N., & Omicini, A. (2010). An Agent-based Model for the Pattern Formation in Drosophila Melanogaster. In *Proceedings of the 12th International Conference on the Synthesis and Simulation of Living Systems*. The MIT Press.

Montagna, S., Omicini, A., & Pianini, D. (2016). Extending the Gillespie's stochastic simulation algorithm for integrating discrete-event and multi-agent based simulation. In B. Gaudou & J. S. Sichman (Eds.), Multi-Agent Based Simulation XVI (Vol. 9568, pp. 3–18). Springer. 1 doi:10.1007/978-3-319-31447-1_1

Montagna, S., Pianini, D., & Viroli, M. (2012). A model for Drosophila Melanogaster development from a single cell to stripe pattern formation. In D. Shin, C.-C. Hung & J. Hong (Eds.), *27th annual ACM Symposium on Applied Computing* (SAC 2012) (pp. 1406–1412). Riva del Garda, Italy: ACM. 10.1145/2245276.2231999

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Montagna, S., Ricci, A., & Omicini, A. (2008). A&A for modelling and engineering simulations in Systems Biology. *International Journal of Agent-Oriented Software Engineering*, *2*(2), 222–245. doi:10.1504/IJAOSE.2008.017316

North, M. J., Collier, N. T., Ozik, J., Tatara, E. R., Macal, C. M., Bragen, M., & Sydelko, P. (2013). Complex adaptive systems modeling with Repast Symphony. *Complex Adaptive Systems Modeling*, *1*(1), 3. doi:10.1186/2194-3206-1-3

Olsen, M. M., & Siegelmann, H. T. (2013). Multiscale agent-based model of tumor angiogenesis. *Procedia Computer Science*, *18*, 1016–1025. doi:10.1016/j.procs.2013.05.267

Omicini, A., & Zambonelli, F. (2004). MAS as complex systems: A view on the role of declarative approaches. In J. A. Leite, A. Omicini, L. Sterling & P. Torroni (Eds.), Declarative Agent Languages and Technologies (Vol. 2990, pp. 1–17). Springer. 1 doi:10.1007/978-3-540-25932-9_1

Pianini, D., Montagna, S., & Viroli, M. (2013). Chemical-oriented simulation of computational systems with Alchemist. *Journal of Simulation*, *7*(S3), 202–215. doi:10.1057/jos.2012.27

Pisarev, A., Poustelnikova, E., Samsonova, M. & Reinitz, J. (2009). FlyEx, the quantitative atlas on segmentation gene expression at cellular resolution. *Nucleic Acids Research, 37*(1), 560–566. doi:10.1093/nar/gkn717

Pogson, M., Smallwood, R., Qwarnstrom, E., & Holcombe, M. (2006). Formal agent-based modelling of intracellular chemical interactions. *Bio Systems*, *85*(1), 37–45. doi:10.1016/j.biosystems.2006.02.004 PMID:16581178

Rouillard, A. D., & Holmes, J. W. (2014). Coupled agent-based and finite-element models for predicting scar structure following myocardial infarction. *Progress in Biophysics and Molecular Biology*, *115*(2–3), 235–243. doi:10.1016/j.pbiomolbio.2014.06.010 PMID:25009995

Setty, Y. (2012). Multi-scale computational modeling of developmental biology. *Bioinformatics (Oxford, England)*, *28*(15), 2022–2028. doi:10.1093/bioinformatics/bts307 PMID:22628522

Slepoy, A., Thompson, A. P., & Plimpton, S. J. (2008). A constant-time kinetic Monte Carlo algorithm for simulation of large biochemical reaction networks. *The Journal of Chemical Physics*, *128*(20), 205101. doi:10.1063/1.2919546 PMID:18513044

Starruß, J., de Back, W., Brusch, L., & Deutsch, A. (2014). Morpheus: A user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics (Oxford, England)*, *30*(9), 1331–1332. doi:10.1093/bioinformatics/btt772 PMID:24443380

Surkova, S., Kosman, D., Kozlov, K., Manu, Myasnikova, E., Samsonova, A. A., ... Reinitz, J. (2008). Characterization of the Drosophila segment determination morphome. *Developmental Biology*, *313*(2), 844–862. doi:10.1016/j.ydbio.2007.10.037 PMID:18067886

Uhrmacher, A. M., Degenring, D., & Zeigler, B. (2005). Discrete event multi-level models for Systems Biology. In C. Priami (Ed.), Transactions on Computational Systems Biology I (Vol. 3380, pp. 66–89). Springer. 6 doi:10.1007/978-3-540-32126-2_6

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Uhrmacher, A. M., & Weyns, D. (Eds.). (2009). *Multi-agent systems: Simulation and applications* (1st ed.). Boca Raton, FL: CRC Press.

Walker, D., Wood, S., Southgate, J., Holcombe, M., & Smallwood, R. (2006). An integrated agent-mathematical model of the effect of intercellular signalling via the epidermal growth factor receptor on cell proliferation. *Journal of Theoretical Biology*, *242*(3), 774–789. doi:10.1016/j.jtbi.2006.04.020 PMID:16765384

Wang, Z., Butner, J. D., Kerketta, R., Cristini, V., & Deisboeck, T. S. (2015). Simulating cancer growth with multiscale agent-based modeling. *Seminars in Cancer Biology*, *30*, 70–78. doi:10.1016/j.semcancer.2014.04.001 PMID:24793698

Weyns, D., Omicini, A., & Odell, J. J. (2007). Environment as a first class abstraction in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, *14*(1), 5–30. doi:10.100710458-006-0012-0

White, T., & Pagurek, B. (1998). Towards multi-swarm problem solving in networks. In *3rd International Conference on Multi Agent Systems (ICMAS '98)* (pp. 333–340). 10.1109/ICMAS.1998.699217

Wilensky, U. (1999). *NetLogo. Center for Connected Learning and Computer-Based Modeling*. Northwestern University. Retrieved from http://ccl.northwestern.edu/netlogo/

Zambonelli, F., & Omicini, A. (2004). Challenges and research directions in agent-oriented software engineering. *Autonomous Agents and Multi-Agent Systems*, *9*(3), 253–283. doi:10.1023/B:AGNT.0000038028.66672.1e

KEY TERMS AND DEFINITIONS

Cellular Potts Model: A lattice model evolved as a time-discrete Markov chain, where the transition probabilities are specified with the help of a *Hamiltonian* or *energy* function.

Cleavage Cycle: A cell cycle occurring during the cleavage stage of embryo development and resulting in the synchronous division of the cells in the early embryo.

Gene: DNA is composed of thousands of genes (the number varies with the organism). A gene contains the set of instructions needed to code a specific protein.

Gene Expression: The process that, beginning with the information contained in a gene and passing through different steps, end up in the synthesis of a functional gene product, that in most of the cases is a protein.

Lattice Model: A physical model where space is defined on a lattice, *i.e.*, on a structure with a regular periodic unit.

mRNA: During the process of gene expression, the first step synthetizes a mRNA molecule that contains all the genetic information stored in the gene and necessary for the next step. From the mRNA cell is then able to code for a specific protein.

Ordinary Differential Equation: One of the most common modelling approaches in Systems Biology is mathematical and adopts the ordinary differential equations. They involve derivative of one or more dependent variables with respect to a single independent variable.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Protein: The basic functional unit of a cell is the protein. Depending on their tridimensional structures, they perform most of the activities inside a living cell.

Reaction Diffusion Systems: They are differential equations with two independent variables that normally account for the change in space and time of molecules.

ENDNOTES

http://www.nature.com/subjects/multicellular-systems

[2] http://apice.unibo.it

[3] http://alchemist.apice.unibo.it

This research was previously published in Multi-Agent-Based Simulations Applied to Biological and Environmental Systems edited by Diana Francisca Adamatti; pages 159-178; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 21
The Role of Artificial Intelligence in Clinical Decision Support Systems and a Classification Framework

Steven Walczak
University of South Florida, USA

ABSTRACT

Clinical decision support systems are meant to improve the quality of decision-making in healthcare. Artificial intelligence is the science of creating intelligent systems that solve complex problems at the level of or better than human experts. Combining artificial intelligence methods into clinical decision support will enable the utilization of large quantities of data to produce relevant decision-making information to practitioners. This article examines various artificial intelligence methodologies and shows how they may be incorporated into clinical decision-making systems. A framework for describing artificial intelligence applications in clinical decision support systems is presented.

INTRODUCTION

As far back as the early 1960's computer technology has been applied to help solve healthcare issues (Saba, Johnson, & Simpson, 1994). Starting in the 1970's, intelligent computer applications to improve healthcare problem solving had arrived (Schwartz, 1970). While Schwartz (1970) incorrectly implied that computers would eventually replace physicians, many of the other forecasts on the impact of computer technology in healthcare, such as improving data availability and improving decision making, have been realized. Nearly four decades later, health information technology professionals continue to advocate for further development of information systems technologies in healthcare to create a nationwide infrastructure for sharing health information to improve patient outcomes and reduce medical costs (Chopra, 2009).

DOI: 10.4018/978-1-7998-1204-3.ch021

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The advent of electronic health records (EHRs) and health information exchanges (HIEs) have certainly provided a means for more rapid and reliable access to patient information to improve the quality of care (Topaz et al., 2013). EHRs are utilized worldwide to improve health delivery systems (Wager, Lee, & Glaser, 2013).

However, EHRs still experience resistance to use from physicians and staff (McAlearney et al., 2015; Wang et al., 2015). Despite improved information flow afforded by EHRs and HIEs, this resistance has resulted in information overload and the need to find better ways to extract reliable information (Huang et al., 2015). For instance, Pickering et al. (2013) demonstrated that physicians use a relatively small amount of information available in EHRs when determining which patients to admit to an ICU, indicating that often too much information is available for decision making and leading to information overload.

Although there is still a long way to go before the ubiquitous acceptance of healthcare information technologies (HIT), healthcare is becoming an ever more information based and information dependent science (Ikram, Ghani, & Abdullah, 2015). Advances in information and communication technologies have led to the availability and generation of vast amounts of data (Esfandiari et al., 2014), especially in clinical medicine (Bellazzi & Zupan, 2008).

Clinical decision support systems, which are defined as "computer systems designed to impact clinician decision making at the point in time that these decisions are made (Berner & Lande, 2007, pg. 3)," are providing mechanisms for delivering clinical information in a usable format to decision makers (Bellazzi & Zupan, 2008). Improving rates of acceptance of clinical decision support systems may be due in part to the concern of physicians and other medical providers about the quantity of information needed to provide high quality health care and to avoid costly and life-threatening errors (Musen, Shahar, & Shortliffe, 2006).

Research has shown that utilizing artificial intelligence alone (Miller & Brown, 2018) or with EHR technology can improve clinical decision support (El-Sappagh & El-Masri, 2014). There is a tremendous need for incorporating artificial intelligence into clinical decision making to improve clinical outcomes (Park & Han, 2018). Artificial intelligence, which is based on the principal of making machines intelligent so that they can perform at expert human levels or better, enables the rapid assimilation of large amounts of highly complex data into multi-criteria decision-making models that can heuristically determine outcomes to clinical and other healthcare related problems (Callahan & Shah, 2018). Applied research in medicine and healthcare is already showing how to use artificial intelligence to improve decision making for detecting and diagnosing diseases or other healthcare issues, planning treatment, allocating critical healthcare resources, and scheduling and monitoring of patients. Pandey and Mishra (2009) document for the field of medicine: 56 knowledge-based expert systems, 33 artificial neural network applications, and 41 genetic algorithm or fuzzy logic-based systems. Other literature review research of artificial neural network, genetic algorithm, and fuzzy set methods applied in medicine found over 1000 reported applications from 1995 through 2007 (Yardimci, 2009) and for intelligent agents solving various treatment and monitoring problems 20 different applications are reported (Isern, Sánchez, & Moreno, 2010). From 1994 through 2014 across 14 specializations of medicine and health care management, Ghaheri et al. (2015) reported on 72 different healthcare information systems applications of genetic algorithms or genetic algorithms used in combination with artificial neural networks. As an indicator of the growth of artificial intelligence research and practice in medicine, a search conducted on EBSCO using the search terms artificial intelligence, artificial neural network, expert system, knowledge based system, machine learning, genetic algorithm, and fuzzy set combined with the search terms medical, medicine,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

doctor, physician, hospital, and healthcare and limited to academic journal publication in English from January 2010 to December 2017 yielded 138,688 results and an identical search on titles and abstracts in PubMed produced 7,670 articles.

Clearly artificial intelligence-based applications research in medical domains is prevalent and continuing to grow. However, knowledge models in many clinical and medical domains are lacking and new knowledge models are still needed (Bohada, 2012). Diagnostic, treatment, and prognosis problems in the field of healthcare are quite complex (Nealon & Moreno, 2003) and consequently lend themselves to artificial intelligence solutions.

HISTORICAL PERSPECTIVE

Interest in applying artificial intelligence to medical domain problems has existed for at least 30 years. The early days of artificial intelligence in medicine primarily saw knowledge-based or expert systems applied to simulate physician expertise in diagnosis and treatment of various medical conditions and automate the reasoning process required in making complex medical decisions. These systems used Socratic-like reasoning methods to provide oracle like results (Miller & Masarie, 1990). One of the early medical expert systems and perhaps the most well-known is MYCIN, introduced in the early 1970's, which diagnosed and planned treatment for bacteriological infections (Shortliffe et al., 1975). Internist-I, developed in the mid-1970's through the early 1980's expanded upon MYCIN by emulating the knowledge of a domain expert across much of internal medicine (Miller, Pople & Myers, 1982). The last example expert system, though there exist many others, is PUFF which was used to diagnose pulmonary disease (Aikins et al., 1983). Many other standalone expert systems were developed following MYCIN through the early 1990's. One of the key features of these expert system approaches is their explanation facilities where they could demonstrate the reasoning involved in reaching and thereby justifying their conclusions.

In the early 1990's focus shifted away from expert system approaches to soft computing approaches including artificial neural networks and genetic algorithms (Yardimci, 2009). The artificial neural network diagnostic approach, similar to the narrow focus of earlier expert systems, focused on solving specific problems. Artificial neural networks enabled the creation of nonparametric, complex, and nonlinear diagnostic models. Unsupervised learning based artificial neural networks were primarily used in medical image processing to facilitate diagnosis, while supervised learning networks were primarily used in clinical diagnostics and other areas such as resource allocation. A full discussion of how expert systems and artificial neural networks work is beyond the scope of this article; however, numerous guides and introductory books exist and the reader is encouraged to peruse these (e.g., see Medsker & Leibowitz, 1993).

The ability of artificial neural networks to capture significant anomalies in medical images with high precision (Egmont-Petersen, de Ridder, & Handels, 2002; Hall et al., 1992; Scott, 1993; Wu et al., 1993) was one gain to the medical profession. Additionally, artificial neural network models of disease diagnosis started to become more accepted, with Reggia (1993) showing 11 examples between 1988 and 1992. The work by Baxt (1990, 1991) helped propel the recognition of artificial neural networks as a legitimate medical diagnostic modeling methodology and his artificial neural network solution for diagnosing acute myocardial infarction was demonstrated to be more accurate than emergency department physicians (Baxt & Skora, 1996).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The twenty-first century brought a focus on combining artificial intelligence methodologies into integrated systems (Pandey & Mishra, 2009). Knowledge-based systems and soft computing systems each have their strengths and also perceived weaknesses, such as the black box nature of artificial neural networks. Combining artificial intelligence techniques into a single hybrid system helps maintain the strengths of a particular artificial intelligence methodology while reducing the effect of any perceived weaknesses, consequently producing even more accurate and dependable decision support systems (Ramesh et al., 2004; Walczak, 2012). Integration of mixed artificial intelligence methods (El-Sappagh & El-Masri, 2014) also helped lead the way for cross platform integration of artificial intelligence into traditional healthcare systems, including EHRs.

ARTIFICIAL INTELLIGENCE IN CLINICAL DECISION SUPPORT SYSTEMS

Traditionally, one of the most significant barriers to utilization of artificial intelligence in healthcare has been lack of physician acceptance (Bhattacherjee & Hikmet, 2007; Walter & Lopez, 2008), a circumstance that also extends to nurses and other healthcare professionals (Khong, Holroyd, & Wang, 2015). Lack of acceptance of artificial intelligence-based HIT has been shown to be an international phenomenon (Gagnon et al., 2014; Hsiao & Chen, 2016). The advent of the 2009 HITECH Act in the United States along with rapid worldwide adoption of EHRs (Wager, Lee, & Glaser, 2013, Appendix C) and additionally the third stage of meaningful use, which requires advanced use of EHR technology to improve patient outcomes (HealthIT.gov, 2015), is opening the door for advanced decision-making applications becoming embedded in traditional healthcare IT.

The implementation of EHRs and the availability of HIEs drastically increases the quantity and quality of information available for decision making (Esfandiari et al., 2014; Goldstein & Shahar, 2016; Patel et al., 2009). Artificial intelligence methods, in particular artificial neural networks, decision trees, and rule-based expert systems are frequently used as datamining approaches for medical information, enabling decreased time spent finding and analyzing relevant evidence and also reducing errors (Obenshain, 2004). Datamining of both structured and unstructured information (e.g., clinical notes) via artificial intelligence methods is having an increasingly important impact on clinical decision making (Robson & Boray, 2016).

The heuristic nature of artificial intelligence enables clinical decision support systems to equal or often surpass the performance of medical personnel (Chappell, Lee, & Taylor, 1992; Kentala et al., 1998; Wolf et al., 2015). The improved decision-making capabilities of artificial intelligence based clinical decision support has several sources. First, expert and other knowledge-based systems using knowledge engineering typically gather the collective experience of multiple experts and encapsulate this expertise into a single system (Hernández-Chan et al., 2016; Wolf et al. 2015). In addition, artificial neural networks and other artificial intelligence soft computing methods such as genetic algorithms employ high powered mathematics to determine complex and sometimes nonlinear associations between independent and dependent clinical variables that are problematic for human experts to discover (Nealon & Moreno, 2003). Artificial intelligence techniques are thus able to apply expert decision-making knowledge incorporating all relevant data regardless of how large and combine these in complex and nonlinear ways to improve the quality of clinical decision making.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A Framework for Classifying AI Applications in CDSS

Artificial intelligence applications in clinical and other medical domains may be classified along a number of different properties. These properties include 1) type of artificial intelligence used in the system, 2) medical specialty or domain in which the application is used, 3) users of the system, 4) beneficiaries of the system, and 5) type of problem being solved (e.g., diagnosis or prognosis). As an example, an artificial neural network application that predicts the transfusion requirements of emergency room (ER) patients (Walczak, 2005) uses the artificial intelligence machine learning method of artificial neural networks, for the medical domain of trauma medicine, used by ER surgeons, benefiting ER patients (as well as hospital administrators), and is a resource allocation type of problem within healthcare. Many artificial intelligence applications in clinical decision support may utilize multiple features within each of these classification types, such a system that uses multiple types of artificial intelligence within a single application like the combination of artificial neural networks and fuzzy logic to perform diagnosis of calcifications in digital mammograms (Verma & Zakos, 2001).

The increased importance of fitting artificial intelligence applications in medicine into the clinical work flow (Patel et al., 2009) emphasizes the need for incorporating the decision maker user into the classification of artificial intelligence based clinical decision support systems.

The first is the artificial intelligence methods being employed within the decision support system: expert system, artificial neural network, genetic algorithm, fuzzy system, intelligent agent, computer vision, natural language processing, support vector machines, other machine learning, or other heuristic methodology. Another useful classification is the target user: physicians, nurses, pharmacists, other healthcare practitioner, healthcare facility, healthcare organization, management, or patient. Within each of these end user types are diverse tasks or goals of the decision support system. Healthcare providers will triage, diagnose, and plan treatment, while facilities and organization may be tasked with making sure sufficient and proper resources are available and patients are tasked with information transfer, scheduling, and health maintenance among other tasks.

The purpose of this article is not to serve as a literature review, but rather to highlight advances in artificial intelligence in clinical decision support and illustrate the capabilities of these systems. Table 1 shows an example of different artificial intelligence methods applied for various clinical end users. As indicated in the literature reviews cited in the Introduction, numerous other examples are available, but beyond the scope of this article.

Another example of using the classification framework is performed using the search terms specified in the Introduction section on Google Scholar. The Google Scholar search was limited to scholarly articles and book chapters appearing from January through April 2018 and was sorted by relevance to the search terms. Analysis of the first 50 articles returned that specify an artificial intelligence system in a medical domain. Abstracts of conference presentations and literature reviews or summaries were then excluded. The analysis is presented in Table 2, which uses the users of the system and the medical specialty or domain for the 2x2 grid and specifies the type of artificial intelligence system used and the type of problem solution within the Table 2 cells.

In addition to the numerous ways in which artificial intelligence supports and improves clinical decision making directly to various providers, artificial intelligence also may contribute

to scientific discovery of new diagnosis and treatment paradigms. Additionally, resource scheduling to complement needs based on clinical decisions may also be accomplished. Walczak (2008) provides

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Sample artificial intelligence applications for various medical user types

	Expert systems	Artificial neural networks	Other artificial intelligence methods
Physicians/ dentists	Expert system to interpret lab results for diagnosis (Osuagwu & Okafor, 2010)	Neural network that predicts occurrence of pulmonary embolism in surgical patients (Walczak, Brimhall, & Lefkowitz, 2006)	Natural language processing system that summarizes longitudinal health data (Goldstein & Shahar, 2016)
Other healthcare professionals*	A fuzzy expert system that evaluates pain level (Padma & Balasubramanie, 2011)	Neural network to model and predict nurse-to-patient communication problems (Bagnasco et al., 2015)	Genetic algorithm to perform constraint satisfaction nurse scheduling (Lin et al., 2015)
Hospitals, clinics & resources	Expert system that helps select locations for new hospitals (Lin, Lee, & Chen, 2010)	Neural network that predicts length of stay (LaFaro et al., 2015)	Natural language processing system to prevent HIPAA violations by detecting PHI (Yang & Garibaldi, 2015)
Patients	Expert system that facilitates interpretation of laboratory test results (Semenov & Kopanitsa, 2016)	Neural network to predict quality of life for oncology patients (Takehira et al., 2011)	Agent based system to provide healthcare over mobile devices (Santos et al., 2016)

* Other healthcare professionals include nurses, nurse practitioners, physician assistants, pharmacists, anesthesiologists, physical and occupational therapists, dental assistants and hygienists, among others.

a framework for utilizing artificial neural networks to evaluate new medical decision-making heuristics and create new heuristics.

Other artificial intelligence methods may also be utilized for evaluation and creation of new heuristics, including: genetic algorithms and other machine learning (ML) methods like case-based reasoning (CBR), hidden Markov models (HMM), and decision trees. Genetic algorithm-based systems have been shown to function well as diagnostic classification systems (Marín et al., 2011). Other machine learning methods include support vector machines (SVM) which have been used to identify all medication information from patient medical records (Patrick & Li, 2010) and recognize EKG heart rhythms (Goh et al., 2006), and Bayesian classifiers which have been used to classify bone fractures and diagnose ischemic heart disease (Kononenko, 2001). These approaches may also be used to validate new clinical guidelines using historic data.

Other Artificial Intelligence Methods in Medicine

Knowledge-based systems and soft computing methods such as artificial neural networks have been highlighted above, but other artificial intelligence research regularly impacts the field of medicine and the world around us. Two such artificial intelligence methods are natural language processing (NLP) and intelligent agent systems.

NLP has been around for a long time and has resulted in some products which many individuals no longer associate with artificial intelligence (Patel et al., 2009). One such tool is the Dragon natural language system that enables clinicians to dictate their notes using a mobile tablet directly to a patient's EHR without need of further transcription and thus time and costs to their medical organization (Peek et al., 2015). As recently as 2006, voice recognition systems were not widely available in hospitals (Ward et al., 2006). Due to the nature of unstructured physician and other clinical personnel communications and notes, NLP is a critical resource for finding, documenting, and reasoning with medical knowledge.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Artificial intelligence in medicine articles published in 2018

	Physician	Other clinicians	Hospital/clinic	Patient
Abdominal (endocrine, excretory, gastrointestinal, hepatic)	diagnostic ML (Feng et al., 2018); diagnostic random forest[1] (López et al., 2018); prognostic fuzzy random forest[1] (Saleh et al., 2018)	prognostic (kidney) ensemble (Kazemi & Mirroshandel, 2018)	N/A	diagnostic (kidney) agent & ANN (Abdelaziz et al., 2018); prognostic (diabetes) ML/CBR (Brown et al., 2018); (diabetes) ANN (Kang, 2018)
Anesthesia/ Surgery	prognostic fuzzy rule-based (Mendez et al., 2018)	N/A	N/A	N/A
Brain	diagnostic ANN (Lacy et al., 2018); prognostic fuzzy system (Garza-Rodríguez et al., 2018); diagnostic SVM & Bayesian (Mumtaz et al., 2018a; Mumtaz et al., 2018b)	N/A	prognostic SVM/ML (Lin et al., 2018)	prognostic SVM and ensemble (Hosseini et al., 2018)
Cellular	diagnostic SVM (Rahman et al., 2018)	N/A	N/A	N/A
Circulatory/ Heart/Blood	diagnostic HMM (Akrivos et al., 2018); diagnostic ML (Gunčar et al., 2018); prognostic ML (Kalscheur et al., 2018)	N/A	diagnostic SVM & ANN (Hu et al., 2018)	prognostic (malaria) Bayesian (Haddawy et al., 2018)
Disease (non-specific)	N/A	N/A	N/A	prognostic (infections) KBS[2]/ML (Shen et al., 2018)
Elderly/Aging	N/A	N/A	prognostic (falls) ANN (Beauchet et al., 2018)	N/A
Oncology	prognostic ANN (Walczak & Velanovich, 2018); diagnostic ML (Xiao et al., 2018)	prognostic ANN (Lee et al., 2018)	N/A	N/A
Opthamology/ Eye	diagnostic ANN ensemble (Orlando et al., 2018); diagnostic ML (Samant & Agarwal, 2018)	N/A	N/A	N/A
Resource planning	N/A	NLP/text analytics for note extraction (Moradi & Ghadiri, 2018)	Markov model for patient scheduling (Li et al., 2018); genetic algorithm for ER planning (Yousefi et al., 2018)	N/A
Respiratory/ Lung	diagnostic ANN (Tajbakhsh & Suzuki, 2018)	N/A	N/A	N/A
Trauma	diagnostic Bayesian (Schetinin et al., 2018)	N/A	GA/ANN ensemble for patient scheduling (Jiang et al., 2018)	diagnostic ML (Nijeweme-d'Hollosy et al., 2018)

NLP may be used to examine and determine applicable clinical guidelines to enable physicians to adhere to existing guidelines. One example is a natural language based clinical decision support system which outperformed domain experts in identifying optimal surveillance periods, according to the guidelines, for colon cancer Imler, Morea, & Imperiale, 2014). Another example is an NLP system embedded

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in an EHR that automatically identified heart disease risk factors from the clinical notes of office visits (Khalifa & Meystre, 2015). Other examples of NLP based clinical decision support systems may be found in Demner-Fushman, Chapman, and McDonald (2009).

Intelligent agents have a long history in the field of medicine as well as more general use. Many people already take advantage of intelligent agents whenever they perform a web search. Early agent-based systems in medicine performed patient appointment reminders and scheduling (Silverman, Andonyadis, & Morales, 1998). Agents enable rapid and efficient search of medical data (Isern, Sánchez, & Moreno, 2010), which is frequently spread across numerous databases, the EHR, and other sources. Recently agents have been employed to deliver treatment planning and monitoring over mobile devices (Chan, Ray, & Parameswaran, 2008; Horan & Schooley, 2007; Santos et al., 2016). The numerous sources and distributed nature of clinical information makes intelligent agents an ideal methodology for working with this type of data.

CHALLENGES AND ROADBLOCKS TO ADOPTION OF ARTIFICIAL INTELLIGENCE IN MEDICINE

As indicated previously the largest challenge for artificial intelligence in clinical decision support systems is acceptance and adoption by physicians. The underlying reasons are often not technological in nature, but stem from organizational, monetary, or cultural reasons (Patel et al., 2009). Additional reasons for the non-acceptance of artificial intelligence-based decision support systems include interruption of workflow, dislike of template-like treatment plans (especially from clinical decision support systems), and wariness of administrative oversight and intervention enabled by HIT systems (Romanow, Cho, & Straub, 2012). One path for improving acceptance and consequent utilization of artificial intelligence-based decision support in clinical systems is the widespread integration of these techniques into existing HIT systems, such as EHRs, which will help alleviate workflow concerns. Other methods to improve acceptance fall into traditional user interface design through creation of simple and easy to understand interfaces to increase perceived ease of use by healthcare professionals. Even when accepted, physicians and other clinicians may feel that artificial intelligence enhanced EHRs have other negative impacts such as reduction in desired social rituals and perceived increase in burnout from alert overloads (Verghese, Shah, & Harrington, 2018).

Heuristic problem solving does save time when datamining medical data sources. However, the heuristic nature of most of artificial intelligence in trying to mimic and surpass human expertise does not always produce the optimal answer. Therefore, it is incumbent on artificial intelligence researchers to build in safeguards and remind medical personnel that the clinical recommendations may require further analysis. The users of artificial intelligence based clinical decision support systems must be educated to understand that these systems can and do make errors, even if errors are less frequent than their human counterparts. For example, an oncology physician may make a correct diagnosis 90 percent of the time and decides to use an artificial intelligence clinical decision support tool which has been shown to deliver accurate diagnosis 95 percent of the time (a 5 percent overall improvement), but this still implies a 5 percent error of misdiagnoses.

While one of the advantages for using expert systems is their explanation facility, informing users of the reasoning employed to make a decision, soft computing artificial intelligence methods such as artificial neural networks and genetic algorithms operate more as a black box. While it is possible to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

generate rules from these black boxes (Hayashi, 2013) and consequent explanations, the process is not commonly used. As indicated above, this doesn't mean that these artificial intelligence methods should be ignored, but rather that discretion is required in applying them to obtain optimal clinical outcomes. The black box nature may also be better accepted by medical personnel as explanation facilities could cause an interruption to the workflow which in turn would lead to non-acceptance of the artificial intelligence clinical decision support tools.

When MYCIN (Shortliffe et al., 1975) was first introduced, artificial intelligence research demonstrated that similar results could be obtained utilizing other artificial intelligence-based methods, such as Bayesian reasoning (Adams, 1976). Yet et al. (2014) point out that simply building models that fit data well is not as important as providing high quality information to healthcare decision makers, both care givers and recipients, to enable optimal healthcare outcomes for the patients. Thus, new research in clinical decision support must keep in mind the desired end goals of improving decision making and the quality of health outcomes. Therefore, clinical artificial intelligence decision support systems must satisfy several criteria (related to technology acceptance): high quality recommendations that fit within existing workflow (perceived usefulness) and results that are easy to interpret and understand (perceived ease of use).

It is critical for artificial intelligence in medicine researchers and application designers to focus on clinical workflow. Additionally, making decision support systems easy to understand and navigate will also increase utilization. Working with the targeted healthcare users of the artificial intelligence clinical decision support is a necessity that cannot be overlooked if these systems are to be used and ultimately improve healthcare processes and outcomes.

CONCLUSION

Clinical information systems and particularly clinical decision support systems are leading the way in improving access to clinical data, improving clinical decision making, and improving the quality of care delivered to patients (Ward et al., 2006). This article has shown how artificial intelligence research and applications serve to further improve decision making in the clinical process from triage to diagnosis to treatment planning.

Artificial intelligence soft computing methods are widely used to perform medical data mining. In addition to acquiring, mining, and organizing medical data, artificial intelligence applications may also serve as a knowledge repository by capturing and preserving clinical expertise and other clinical decision-making protocols or guidelines. This expertise may then be shared and utilized by other clinical personnel as well as directly consumed by patients.

Patients are benefiting from artificial intelligence research applications to become better informed health consumers. Intelligent agents serve to locate reliable health information to assist with patient decision making (Santos et al., 2016) and also to assist with monitoring and following treatment plans (Silverman, Andonyadis, & Morales, 1998). Expert systems and natural language processing assist patients in understanding medical jargon including laboratory tests (Semenov & Kopanitsa, 2016). Future artificial intelligence research should continue to provide applications for patients.

Artificial intelligence research in medicine and applied in clinical decision support systems will continue to improve the efficiency and effectiveness of care while simultaneously helping to reduce healthcare costs. Applications of artificial intelligence decision making must respect the clinical work-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

flow. One way to accomplish this is to continue research on embedding artificial intelligence applications into existing clinical decision support systems already used by physicians and other medical staff (El-Sappagh & El-Masri, 2014), such as EHRs.

Ideally, artificial intelligence research in clinical decision making is focused on making sure that the correct diagnosis is always made and the optimal treatment plan for each patient is discovered. Artificial intelligence in medicine research however needs support and funding to accomplish these goals. Governments and other investors need to continue funding and promoting artificial intelligence research so that the quality of clinical decision making may continue to be improved for the timely, cost effective, and high-quality delivery of care in clinical settings.

REFERENCES

Abdelaziz, A., Elhoseny, M., Salama, A. S., & Riad, A. M. (2018). A machine learning model for improving healthcare services on cloud computing environment. *Measurement*, *119*, 117–128. doi:10.1016/j.measurement.2018.01.022

Adams, J. B. (1976). A probability model of medical reasoning and the MYCIN model. *Mathematical Biosciences*, *32*(1), 177–186. doi:10.1016/0025-5564(76)90064-X

Aikins, J. S., Kunz, J. C., Shortliffe, E. H., & Fallat, R. J. (1983). PUFF: An expert system for interpretation of pulmonary function data. *Computers and Biomedical Research, an International Journal*, *16*(3), 199–208. doi:10.1016/0010-4809(83)90021-6 PMID:6347509

Akrivos, E., Papaioannou, V., Maglaveras, N., & Chouvarda, I. (2018). Prediction of Cardiac Arrest in Intensive Care Patients Through Machine Learning. In *Precision Medicine Powered by pHealth and Connected Health* (pp. 25–29). Singapore: Springer. doi:10.1007/978-981-10-7419-6_5

Bagnasco, A., Siri, A., Aleo, G., Rocco, G., & Sasso, L. (2015). Applying artificial neural networks to predict communication risks in the emergency department. *Journal of Advanced Nursing*, *71*(10), 2293–2304. doi:10.1111/jan.12691 PMID:26010174

Baxt, W. G. (1990). Use of an artificial neural network for data analysis in clinical decision-making: The diagnosis of acute coronary occlusion. *Neural Computation*, *2*(4), 480–489. doi:10.1162/neco.1990.2.4.480

Baxt, W. G. (1991). Use of an artificial neural network for the diagnosis of myocardial infarction. *Annals of Internal Medicine*, *115*(11), 843–848. doi:10.7326/0003-4819-115-11-843 PMID:1952470

Baxt, W. G., & Skora, J. (1996). Prospective validation of artificial neural network trained to identify acute myocardial infarction. *Lancet*, *347*(8993), 12–15. doi:10.1016/S0140-6736(96)91555-X PMID:8531540

Beauchet, O., Noublanche, F., Simon, R., Sekhon, H., Chabot, J., Levinoff, E. J., ... Launay, C. P. (2018). Falls risk prediction for older inpatients in acute care medical wards: Is there an interest to combine an early nurse assessment and the artificial neural network analysis? *The Journal of Nutrition, Health & Aging*, *22*(1), 131–137. doi:10.100712603-017-0950-z PMID:29300432

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bellazzi, R., & Zupan, B. (2008). Predictive data mining in clinical medicine: Current issues and guidelines. *International Journal of Medical Informatics*, *77*(2), 81–97. doi:10.1016/j.ijmedinf.2006.11.006 PMID:17188928

Berner, E. S., & La Lande, T. J. (2007). Overview of Clinical Decision Support Systems. In E. A. Berner (Ed.), *Clinical Decision Support Systems Theory and Practice* (2nd ed., pp. 3–22). New York: Springer. doi:10.1007/978-0-387-38319-4_1

Bhattacherjee, A., & Hikmet, N. (2007). Physicians' Resistance toward Healthcare Information Technology: A Theoretical Model and Empirical Test. *European Journal of Information Systems*, *16*(6), 725–737. doi:10.1057/palgrave.ejis.3000717

Bohada, J. A. (2012). Automatic Production and Integration of Knowledge to the Support of the Decision and Planning Activities in Medical-Clinical Diagnosis, Treatment and Prognosis [Doctoral dissertation].

Brown, D., Aldea, A., Harrison, R., Martin, C., & Bayley, I. (2018). Temporal case-based reasoning for type 1 diabetes mellitus bolus insulin decision support. *Artificial Intelligence in Medicine*, *85*, 28–42. doi:10.1016/j.artmed.2017.09.007 PMID:28986108

Callahan, A., & Shah, N. H. (2018). Machine Learning in Healthcare. In Key Advances in Clinical Informatics (pp. 279-291).

Chan, V., Ray, P., & Parameswaran, N. (2008). Mobile e-Health monitoring: An agent-based approach. *IET Communications*, *2*(2), 223–230. doi:10.1049/iet-com:20060646

Chappell, G. J., Lee, J., & Taylor, J. G. (1992). A Review of Medical Diagnostic Applications of Neural Networks. In J. G. Taylor (Ed.), *Neural Network Applications* (pp. 23–26). London: Springer. doi:10.1007/978-1-4471-2003-2_3

Chopra, A. (2009). [Keynote Address]. In *CONNECT Seminar 2009*, Washington DC. Retrieved from https://www.youtube.com/watch?v=xsTdroooSV8

Demner-Fushman, D., Chapman, W. W., & McDonald, C. J. (2009). What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics*, *42*(5), 760–772. doi:10.1016/j.jbi.2009.08.007 PMID:19683066

Egmont-Petersen, M., de Ridder, D., & Handels, H. (2002). Image processing with neural networks—a review. *Pattern Recognition*, *35*(10), 2279–2301. doi:10.1016/S0031-3203(01)00178-9

El-Sappagh, S. H., & El-Masri, S. (2014). A distributed clinical decision support system architecture. *Journal of King Saud University-Computer and Information Sciences*, *26*(1), 69–78. doi:10.1016/j.jksuci.2013.03.005

Esfandiari, N., Babavalian, M. R., Moghadam, A. M. E., & Tabar, V. K. (2014). Knowledge Discovery in Medicine: Current Issue and Future Trend. *Expert Systems with Applications*, *41*(9), 4434–4463. doi:10.1016/j.eswa.2014.01.011

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Feng, Z., Rong, P., Cao, P., Zhou, Q., Zhu, W., Yan, Z., ... Wang, W. (2018). Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. *European Radiology*, *28*(4), 1625–1633. doi:10.100700330-017-5118-z PMID:29134348

Gagnon, M. P., Ghandour, E. K., Talla, P. K., Simonyan, D., Godin, G., Labrecque, M., ... Rousseau, M. (2014). Electronic health record acceptance by physicians: Testing an integrated theoretical model. *Journal of Biomedical Informatics*, *48*, 17–27. doi:10.1016/j.jbi.2013.10.010 PMID:24184678

Garza-Rodríguez, A., Sánchez-Pérez, L. A., Ornelas-Vences, C., & Ehrenberg-Inzunza, M. (2018). Pronation and supination analysis based on biomechanical signals from Parkinson's disease patients. *Artificial Intelligence in Medicine*, *84*, 7–22. doi:10.1016/j.artmed.2017.10.001 PMID:29042162

Ghaheri, A., Shoar, S., Naderan, M., & Hoseini, S. S. (2015). The Applications of Genetic Algorithms in Medicine. *Oman Medical Journal*, *30*(6), 406–416. doi:10.5001/omj.2015.82 PMID:26676060

Goh, K. W., Lavanya, J., Tan, E. K., Soh, C. B., & Kim, Y. (2006, January). A PDA-based ECG beat detector for home cardiac care. In *27th Annual International Conference of the Engineering in Medicine and Biology Society IEEE-EMBS 2005* (pp. 375-378). IEEE.

Goldstein, A., & Shahar, Y. (2016). An automated knowledge-based textual summarization system for longitudinal, multivariate clinical data. *Journal of Biomedical Informatics*, *61*, 159–175. doi:10.1016/j.jbi.2016.03.022 PMID:27039119

Gunčar, G., Kukar, M., Notar, M., Brvar, M., Černelč, P., Notar, M., & Notar, M. (2018). An application of machine learning to haematological diagnosis. *Scientific Reports*, *8*(1), 411. doi:10.103841598-017-18564-8 PMID:29323142

Haddawy, P., Hasan, A. I., Kasantikul, R., Lawpoolsri, S., Sa-angchai, P., Kaewkungwal, J., & Singhasivanon, P. (2018). Spatiotemporal Bayesian networks for malaria prediction. *Artificial Intelligence in Medicine*, *84*, 127–138. doi:10.1016/j.artmed.2017.12.002 PMID:29241658

Hall, L. O., Bensaid, A. M., Clarke, L. P., Velthuizen, R. P., Silbiger, M. S., & Bezdek, J. C. (1992). A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. *IEEE Transactions on Neural Networks*, *3*(5), 672–682. doi:10.1109/72.159057 PMID:18276467

Hayashi, Y. (2013). Neural network rule extraction by a new ensemble concept and its theoretical and historical background: A review. *International Journal of Computational Intelligence and Applications, 12*(4), 1-22.

Health, I. T. gov. (2015). Health IT Regulations – Meaningful Use Regulations. Retrieved from https://www.healthit.gov/policy-researchers-implementers/meaningful-use-regulations

Hernández-Chan, G. S., Ceh-Varela, E. E., Sanchez-Cervantes, J. L., Villanueva-Escalante, M., Rodríguez-González, A., & Pérez-Gallardo, Y. (2016). Collective intelligence in medical diagnosis systems: A case study. *Computers in Biology and Medicine*, *74*, 45–53. doi:10.1016/j.compbiomed.2016.04.016 PMID:27177267

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Horan, T. A., & Schooley, B. L. (2007). Time-critical information services. *Communications of the ACM*, *50*(3), 73–78. doi:10.1145/1226736.1226738

Hosseini, M. P., Pompili, D., Elisevich, K., & Soltanian-Zadeh, H. (2018). Random ensemble learning for EEG classification. *Artificial Intelligence in Medicine*, *84*, 146–158. doi:10.1016/j.artmed.2017.12.004 PMID:29306539

Hsiao, J. L., & Chen, R. F. (2016). Critical factors influencing physicians' intention to use computerized clinical practice guidelines: An integrative model of activity theory and the technology acceptance model. *BMC Medical Informatics and Decision Making*, *16*(1), 1–15. PMID:26772169

Hu, Y., Duan, K., Zhang, Y., Hossain, M. S., Rahman, S. M. M., & Alelaiwi, A. (2018). Simultaneously aided diagnosis model for outpatient departments via healthcare big data analytics. *Multimedia Tools and Applications*, *77*(3), 3729–3743. doi:10.100711042-016-3719-1

Huang, C. W., Lu, R., Iqbal, U., Lin, S. H., Nguyen, P. A., Yang, H. C., ... Jian, W. S. (2014). A richly interactive exploratory data analysis and visualization tool using electronic medical records. *BMC Medical Informatics and Decision Making*, *15*(1), 1–14. PMID:26563282

Ikram, R. R. R., Ghani, M. K. A., & Abdullah, N. (2015). An analysis of application of health informatics in Traditional Medicine: A review of four Traditional Medicine Systems. *International Journal of Medical Informatics*, *84*(11), 988–996. doi:10.1016/j.ijmedinf.2015.05.007 PMID:26160148

Imler, T. D., Morea, J., & Imperiale, T. F. (2014). Clinical decision support with natural language processing facilitates determination of colonoscopy surveillance intervals. *Clinical Gastroenterology and Hepatology*, *12*(7), 1130–1136. doi:10.1016/j.cgh.2013.11.025 PMID:24316106

Isern, D., Sánchez, D., & Moreno, A. (2010). Agents applied in health care: A review. *International Journal of Medical Informatics*, *79*(3), 145–166. doi:10.1016/j.ijmedinf.2010.01.003 PMID:20129820

Jiang, S., Chin, K. S., & Tsui, K. L. (2018). A universal deep learning approach for modeling the flow of patients under different severities. *Computer Methods and Programs in Biomedicine*, *154*, 191–203. doi:10.1016/j.cmpb.2017.11.003 PMID:29249343

Kalscheur, M. M., Kipp, R. T., Tattersall, M. C., Mei, C., Buhr, K. A., DeMets, D. L., ... Page, C. D. (2018). Machine Learning Algorithm Predicts Cardiac Resynchronization Therapy Outcomes: Lessons from the COMPANION Trial. *Circulation: Arrhythmia and Electrophysiology*, *11*(1), e005499. doi:10.1161/CIRCEP.117.005499 PMID:29326129

Kang, S. (2018). Personalized prediction of drug efficacy for diabetes treatment via patient-level sequential modeling with neural networks. *Artificial Intelligence in Medicine*, *85*, 1–6. doi:10.1016/j.artmed.2018.02.004 PMID:29482961

Kazemi, Y., & Mirroshandel, S. A. (2018). A novel method for predicting kidney stone type using ensemble learning. *Artificial Intelligence in Medicine*, *84*, 117–126. doi:10.1016/j.artmed.2017.12.001 PMID:29241659

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kentala, E., Juhola, M., Auramo, Y., & Pyykkö, I. (1998). Comparison between diagnoses of human experts and a neurotologic expert system. *The Annals of Otology, Rhinology, and Laryngology*, *107*(2), 135–140. doi:10.1177/000348949810700209 PMID:9486908

Khalifa, A., & Meystre, S. (2015). Adapting existing natural language processing resources for cardiovascular risk factors identification in clinical notes. *Journal of Biomedical Informatics*, *58*, S128–S132. doi:10.1016/j.jbi.2015.08.002 PMID:26318122

Khong, P. C. B., Holroyd, E., & Wang, W. (2015). A Critical Review of the Theoretical Frameworks and the Conceptual Factors in the Adoption of Clinical Decision Support Systems. *Computers, Informatics, Nursing*, *33*(12), 555–570. doi:10.1097/CIN.0000000000000196 PMID:26535769

Kononenko, I. (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine*, *23*(1), 89–109. doi:10.1016/S0933-3657(01)00077-X PMID:11470218

Lacy, S. E., Smith, S. L., & Lones, M. A. (2018). Using echo state networks for classification: A case study in Parkinson's disease diagnosis. *Artificial Intelligence in Medicine*, *86*, 53–59. doi:10.1016/j.artmed.2018.02.002 PMID:29475631

LaFaro, R. J., Pothula, S., Kubal, K. P., Inchiosa, M. E., Pothula, V. M., Yuan, S. C., ... Perline, R. (2015). Neural Network Prediction of ICU Length of Stay Following Cardiac Surgery Based on Pre-Incision Variables. *PLoS One*, *10*(12), e0145395. doi:10.1371/journal.pone.0145395 PMID:26710254

Lee, S. I., Celik, S., Logsdon, B. A., Lundberg, S. M., Martins, T. J., Oehler, V. G., ... Saxena, A. (2018). A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nature Communications*, *9*(1), 42. doi:10.103841467-017-02465-5 PMID:29298978

Li, X., Wang, J., & Fung, R. Y. (2018). Approximate dynamic programming approaches for appointment scheduling with patient preferences. *Artificial Intelligence in Medicine*, *85*, 16–25. doi:10.1016/j.artmed.2018.02.001 PMID:29482960

Lin, C. C., Kang, J. R., Chiang, D. J., & Chen, C. L. (2015). Nurse scheduling with joint normalized shift and day-off preference satisfaction using a genetic algorithm with immigrant scheme. *International Journal of Distributed Sensor Networks*, *2015*, 1–10. doi:10.1155/2015/674591

Lin, C. T., Lee, C., & Chen, Z. J. (2010). An Expert System Approach to Medical Region Selection for a New Hospital Using Data Envelopment Analysis. *I-Business*, *2*(2), 128.

Lin, W. Y., Chen, C. H., Tseng, Y. J., Tsai, Y. T., Chang, C. Y., Wang, H. Y., & Chen, C. K. (2018). Predicting Post-Stroke Activities of Daily Living Through a Machine Learning-Based Approach on Initiating Rehabilitation. *International Journal of Medical Informatics*, *111*, 159–164. doi:10.1016/j.ijmedinf.2018.01.002 PMID:29425627

López, B., Torrent-Fontbona, F., Viñas, R., & Fernández-Real, J. M. (2018). Single Nucleotide Polymorphism relevance learning with Random Forests for Type 2 diabetes risk prediction. *Artificial Intelligence in Medicine*, *85*, 43–49. doi:10.1016/j.artmed.2017.09.005 PMID:28943335

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Marín, O., Pérez, I., Ruiz, D., Soriano, A., & García, J. D. (2011). Neural networks versus genetic algorithms as medical classifiers. In J. M. Ferrandez, J. R. A. Sanchez, F. de la Paz, & F. J. Toledo (Eds.), *Foundations on Natural and Artificial Computation Part 1* (pp. 393–400). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-21344-1_41

McAlearney, A. S., Hefner, J. L., Sieck, C. J., & Huerta, T. R. (2015). The Journey through Grief: Insights from a Qualitative Study of Electronic Health Record Implementation. *Health Services Research*, *50*(2), 462–488. doi:10.1111/1475-6773.12227 PMID:25219627

Medsker, L. R., & Liebowitz, J. (1993). *Design and Development of Expert Systems and Neutral Networks*. Upper Saddle River, NJ: Prentice Hall.

Mendez, J. A., Leon, A., Marrero, A., Gonzalez-Cava, J. M., Reboso, J. A., Estevez, J. I., & Gomez-Gonzalez, J. F. (2018). Improving the anesthetic process by a fuzzy rule based medical decision system. *Artificial Intelligence in Medicine*, *84*, 159–170. doi:10.1016/j.artmed.2017.12.005 PMID:29310966

Miller, D. D., & Brown, E. W. (2018). Artificial Intelligence in Medical Practice: The Question to the Answer? *The American Journal of Medicine*, *131*(2), 129–133. doi:10.1016/j.amjmed.2017.10.035 PMID:29126825

Miller, R. A., & Masari, F. E. Jr. (1990). The demise of the "Greek Oracle" model for medical diagnostic systems. *Methods of Information in Medicine*, *29*(1), 1–2. doi:10.1055-0038-1634767 PMID:2407929

Miller, R. A., Pople, H. E. Jr, & Myers, J. D. (1982). Internist-I, an Experimental Computer-Based Diagnostic Consultant for General Internal Medicine. *The New England Journal of Medicine*, *307*(8), 468–476. doi:10.1056/NEJM198208193070803 PMID:7048091

Moradi, M., & Ghadiri, N. (2018). Different approaches for identifying important concepts in probabilistic biomedical text summarization. *Artificial Intelligence in Medicine*, *84*, 101–116. doi:10.1016/j.artmed.2017.11.004 PMID:29208328

Mumtaz, W., Ali, S. S. A., Yasin, M. A. M., & Malik, A. S. (2018a). A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD). *Medical & Biological Engineering & Computing*, *56*(2), 233–246. doi:10.100711517-017-1685-z PMID:28702811

Mumtaz, W., Kamel, N., Ali, S. S. A., & Malik, A. S. (2018b). An EEG-based functional connectivity measure for automatic detection of alcohol use disorder. *Artificial Intelligence in Medicine*, *84*, 79–89. doi:10.1016/j.artmed.2017.11.002 PMID:29169647

Musen, M. A., Shahar, Y., & Shortliffe, E. H. (2006). Clinical Decision-Support Systems. In *Biomedical Informatics* (pp. 698–736). New York: Springer. doi:10.1007/0-387-36278-9_20

Nealon, J., & Moreno, A. (2003). Agent-based applications in health care. In A. Moreno & J. L. Nealon (Eds.), *Applications of Software Agent Technology in The Health Care Domain* (pp. 3–18). Basel: Birkhäuser. doi:10.1007/978-3-0348-7976-7_2

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nijeweme-d'Hollosy, W. O., van Velsen, L., Poel, M., Groothuis-Oudshoorn, C. G., Soer, R., & Hermens, H. (2018). Evaluation of three machine learning models for self-referral decision support on low back pain in primary care. *International Journal of Medical Informatics*, *110*, 31–41. doi:10.1016/j.ijmedinf.2017.11.010 PMID:29331253

Obenshain, M. K. (2004). Application of Data Mining Techniques to Healthcare Data. *Infection Control and Hospital Epidemiology*, *25*(8), 690–695. doi:10.1086/502460 PMID:15357163

Orlando, J. I., Prokofyeva, E., del Fresno, M., & Blaschko, M. B. (2018). An ensemble deep learning based approach for red lesion detection in fundus images. *Computer Methods and Programs in Biomedicine*, *153*, 115–127. doi:10.1016/j.cmpb.2017.10.017 PMID:29157445

Osuagwu, C. C., & Okafor, E. C. (2010). Framework for eliciting knowledge for a medical laboratory diagnostic expert system. *Expert Systems with Applications*, *37*(7), 5009–5016. doi:10.1016/j.eswa.2009.12.012

Padma, T., & Balasubramanie, P. (2011). Domain experts' knowledge-based intelligent decision support system in occupational shoulder and neck pain therapy. *Applied Soft Computing*, *11*(2), 1762–1769. doi:10.1016/j.asoc.2010.05.019

Pandey, B., & Mishra, R. B. (2009). Knowledge and intelligent computing system in medicine. *Computers in Biology and Medicine*, *39*(3), 215–230. doi:10.1016/j.compbiomed.2008.12.008 PMID:19201398

Park, S. H., & Han, K. (2018). Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology*, *286*(3), 800–809. doi:10.1148/radiol.2017171920 PMID:29309734

Patel, V. L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R., & Abu-Hanna, A. (2009). The coming of age of artificial intelligence in medicine. *Artificial Intelligence in Medicine*, *46*(1), 5–17. doi:10.1016/j.artmed.2008.07.017 PMID:18790621

Patrick, J., & Li, M. (2010). High accuracy information extraction of medication information from clinical notes: 2009 i2b2 medication extraction challenge. *Journal of the American Medical Informatics Association*, *17*(5), 524–527. doi:10.1136/jamia.2010.003939 PMID:20819856

Peek, N., Combi, C., Marin, R., & Bellazzi, R. (2015). Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes. *Artificial Intelligence in Medicine*, *65*(1), 61–73. doi:10.1016/j.artmed.2015.07.003 PMID:26265491

Pickering, B. W., Gajic, O., Ahmed, A., Herasevich, V., & Keegan, M. T. (2013). Data Utilization for Medical Decision Making at the Time of Patient Admission to ICU. *Critical Care Medicine*, *41*(6), 1502–1510. doi:10.1097/CCM.0b013e318287f0c0 PMID:23528804

Rahman, M. S., Rahman, M. K., Kaykobad, M., & Rahman, M. S. (2018). isGPT: An optimized model to identify sub-Golgi protein types using SVM and Random Forest based feature selection. *Artificial Intelligence in Medicine*, *84*, 90–100. doi:10.1016/j.artmed.2017.11.003 PMID:29183738

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ramesh, A. N., Kambhampati, C., Monson, J. R. T., & Drew, P. J. (2004). Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*, *86*(5), 334–338. doi:10.1308/147870804290 PMID:15333167

Reggia, J. A. (1993). Neural computation in medicine. *Artificial Intelligence in Medicine*, *5*(2), 143–157. doi:10.1016/0933-3657(93)90014-T PMID:8358491

Robson, B., & Boray, S. (2016). Data-mining to build a knowledge representation store for clinical decision support. Studies on curation and validation based on machine performance in multiple choice medical licensing examinations. *Computers in Biology and Medicine*, *73*, 71–93. doi:10.1016/j.compbiomed.2016.02.010 PMID:27089305

Romanow, D., Cho, S., & Straub, D. (2012). Editor's Comments: Riding the Wave: Past Trends and Future Directions for Health IT Research. *Management Information Systems Quarterly*, *36*(3), iii–A18.

Saba, V. K., Johnson, J. E., & Simpson, R. L. (1994). *Computers in Nursing Management*. Washington, DC: ANA.

Saleh, E., Błaszczyński, J., Moreno, A., Valls, A., Romero-Aroca, P., de la Riva-Fernández, S., & Słowiński, R. (2018). Learning ensemble classifiers for diabetic retinopathy assessment. *Artificial Intelligence in Medicine*, *85*, 50–63. doi:10.1016/j.artmed.2017.09.006 PMID:28993124

Samant, P., & Agarwal, R. (2018). Machine learning techniques for medical diagnosis of diabetes using iris images. *Computer Methods and Programs in Biomedicine*, *157*, 121–128. doi:10.1016/j.cmpb.2018.01.004 PMID:29477420

Santos, J., Rodrigues, J. J., Silva, B. M., Casal, J., Saleem, K., & Denisov, V. (2016). An IoT-based mobile gateway for intelligent personal assistants on mobile health environments. *Journal of Network and Computer Applications*, *71*, 194–204. Retrieved from http://www.sciencedirect.com/science/article/pii/S1084804516300273. doi:10.1016/j.jnca.2016.03.014

Schetinin, V., Jakaite, L., & Krzanowski, W. (2018). Bayesian averaging over Decision Tree models for trauma severity scoring. *Artificial Intelligence in Medicine*, *84*, 139–145. doi:10.1016/j.artmed.2017.12.003 PMID:29275896

Schwartz, W. B. (1970). Medicine and the Computer: The Promise and Problems of Change. *The New England Journal of Medicine*, *283*(23), 1257–1264. doi:10.1056/NEJM197012032832305 PMID:4920342

Scott, R. (1993). Artificial Intelligence: Its Use in Medical Diagnosis. *Journal of Nuclear Medicine*, *34*(3), 510–514. PMID:8441047

Semenov, I., & Kopanitsa, G. (2016). Implementation of a Decision Support System for Interpretation of Laboratory Tests for Patients. In J. Hofdijk et al. (Eds.), *Transforming Healthcare with the Internet of Things* (pp. 79–83). Amsterdam, The Netherlands: IOS Press.

Shen, Y., Yuan, K., Chen, D., Colloc, J., Yang, M., Li, Y., & Lei, K. (2018). An ontology-driven clinical decision support system (IDDAP) for infectious disease diagnosis and antibiotic prescription. *Artificial Intelligence in Medicine*, *86*, 20–32. doi:10.1016/j.artmed.2018.01.003 PMID:29433958

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shortliffe, E. H., Davis, R., Axline, S. G., Buchanan, B. G., Green, C. C., & Cohen, S. N. (1975). Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Computers and Biomedical Research, an International Journal*, *8*(4), 303–320. doi:10.1016/0010-4809(75)90009-9 PMID:1157471

Silverman, B. G., Andonyadis, C., & Morales, A. (1998). Web-based health care agents; the case of reminders and todos, too (R2Do2). *Artificial Intelligence in Medicine*, *14*(3), 295–316. doi:10.1016/S0933-3657(98)00039-6 PMID:9821519

Tajbakhsh, N., & Suzuki, K. (2018). A Comparative Study of Modern Machine Learning Approaches for Focal Lesion Detection and Classification in Medical Images: BoVW, CNN and MTANN. In *Artificial Intelligence in Decision Support Systems for Diagnosis in Medical Imaging* (pp. 31–58). Cham, Switzerland: Springer. doi:10.1007/978-3-319-68843-5_2

Takehira, R., Murakami, K., Katayama, S., Nishizawa, K., & Yamamura, S. (2011). Artificial Neural Network Modeling of Quality of Life of Cancer Patients: Relationships between Quality of Life Assessments, as Evaluated by Patients, Pharmacists, and Nurses. *International Journal of Biomedical Science; IJBS*, *7*(4), 255–262. PMID:23675244

Topaz, M., Rao, A., Creber, R. M., & Bowles, K. H. (2013). Educating clinicians on new elements incorporated into the electronic health record: Theories, evidence, and one educational project. *Computers, informatics, nursing*, *31*(8), 375. PMID:23774448

Verghese, A., Shah, N. H., & Harrington, R. A. (2018). What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. *Journal of the American Medical Association*, *319*(1), 19–20. doi:10.1001/jama.2017.19198 PMID:29261830

Verma, B., & Zakos, J. (2001). A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. *IEEE Transactions on Information Technology in Biomedicine*, *5*(1), 46–54. doi:10.1109/4233.908389 PMID:11300216

Wager, K. A., Lee, F. W., & Glaser, J. P. (2013). *Health Care Information Systems: A Practical Approach for Health Care Management*. San Francisco: Jossey-Bass/Wiley.

Wager, K. A., Lee, F. W., & Glaser, J. P. (2013). *Health Care Information Systems: A Practical Approach for Health Care Management* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Walczak, S. (2005). Artificial Neural Network Medical Decision Support Tool: Predicting Transfusion Requirements of ER Patients. *IEEE Transactions on Information Technology in Biomedicine*, *9*(3), 468–474. doi:10.1109/TITB.2005.847510 PMID:16167701

Walczak, S. (2008). Evaluating Medical Decision Making Heuristics and Other Business Heuristics with Neural Networks. In G. Phillips-Wren & N. Ichalkaranje (Eds.), *Intelligent Decision Making: An AI-Based Approach* (pp. 259–287). Berlin: Springer. doi:10.1007/978-3-540-76829-6_10

Walczak, S. (2012). Methodological Triangulation Using Neural Networks for Business Research. *Advances in Artificial Neural Systems*, (517234), 1–12. doi:10.1155/2012/517234

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Walczak, S., Brimhall, B. B., & Lefkowitz, J. B. (2006). Nonparametric decision support systems in medical diagnosis: Modeling pulmonary embolism. *International Journal of Healthcare Information Systems and Informatics*, *1*(2), 65–82. doi:10.4018/jhisi.2006040105

Walczak, S., & Velanovich, V. (2018). Improving prognosis and reducing decision regret for pancreatic cancer treatment using artificial neural networks. *Decision Support Systems*, *106*, 110–118. doi:10.1016/j.dss.2017.12.007

Walter, Z., & Lopez, M. S. (2008). Physician acceptance of information technologies: Role of perceived threat to professional autonomy. *Decision Support Systems*, *46*(1), 206–215. doi:10.1016/j.dss.2008.06.004

Wang, J. Y., Ho, H. Y., Chen, J. D., Chai, S., Tai, C. J., & Chen, Y. F. (2015). Attitudes toward inter-hospital electronic patient record exchange: Discrepancies among physicians, medical record staff, and patients. *BMC Health Services Research*, *15*(1), 1–15. doi:10.118612913-015-0896-y PMID:26162748

Ward, M. M., Jaana, M., Bahensky, J. A., Vartak, S., & Wakefield, D. S. (2006). Clinical Information System Availability and Use in Urban and Rural Hospitals. *Journal of Medical Systems*, *30*(6), 429–438. doi:10.100710916-006-9014-3 PMID:17233155

Wolf, M., Krause, J., Carney, P. A., Bogart, A., & Kurvers, R. H. (2015). Collective Intelligence Meets Medical Decision-Making: The Collective Outperforms the Best Radiologist. *PLoS One*, *10*(8), e0134269. doi:10.1371/journal.pone.0134269 PMID:26267331

Wu, Y., Giger, M. L., Doi, K., Vyborny, C. J., Schmidt, R. A., & Metz, C. E. (1993). Artificial neural networks in mammography: Application to decision making in the diagnosis of breast cancer. *Radiology*, *187*(1), 81–87. doi:10.1148/radiology.187.1.8451441 PMID:8451441

Xiao, Y., Wu, J., Lin, Z., & Zhao, X. (2018). A deep learning-based multi-model ensemble method for cancer prediction. *Computer Methods and Programs in Biomedicine*, *153*, 1–9. doi:10.1016/j.cmpb.2017.09.005 PMID:29157442

Yang, H., & Garibaldi, J. M. (2015). Automatic detection of protected health information from clinic narratives. *Journal of Biomedical Informatics*, *58*, S30–S38. doi:10.1016/j.jbi.2015.06.015 PMID:26231070

Yardimci, A. (2009). Soft computing in medicine. *Applied Soft Computing*, *9*(3), 1029–1043. doi:10.1016/j.asoc.2009.02.003

Yet, B., Perkins, Z., Fenton, N., Tai, N., & Marsh, W. (2014). Not just data: A method for improving prediction with knowledge. *Journal of Biomedical Informatics*, *48*, 28–37. doi:10.1016/j.jbi.2013.10.012 PMID:24189161

Yousefi, M., Yousefi, M., Ferreira, R. P. M., Kim, J. H., & Fogliatto, F. S. (2018). Chaotic genetic algorithm and Adaboost ensemble metamodeling approach for optimum resource planning in emergency departments. *Artificial Intelligence in Medicine*, *84*, 23–33. doi:10.1016/j.artmed.2017.10.002 PMID:29054572

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] Random forest is a collection of decision trees networked together.
[2] KBS = Knowledge-based system or expert system

This research was previously published in the International Journal of Computers in Clinical Practice (IJCCP), 3(2); edited by Athina A. Lazakidou; pages 31-47; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 22
Predictive Modeling of Surgical Site Infections Using Sparse Laboratory Data

Prabhu RV Shankar
University of California Davis, USA

Anupama Kesari
Sri Jayachamarajendra College of Engineering, India

Priya Shalini
Sri Jayachamarajendra College of Engineering, India

N. Kamalashree
Sri Jayachamarajendra College of Engineering, India

Charan Bharadwaj
Sri Jayachamarajendra College of Engineering, India

Nitika Raj
Sri Jayachamarajendra College of Engineering, India

Sowrabha Srinivas
Sri Jayachamarajendra College of Engineering, India

Manu Shivakumar
State University of New York at Buffalo, USA

Anand Raj Ulle
Sri Jayachamarajendra College of Engineering, India

Nagabhushana N. Tagadur
Sri Jayachamarajendra College of Engineering, India

ABSTRACT

As part of a data mining competition, a training and test set of laboratory test data about patients with and without surgical site infection (SSI) were provided. The task was to develop predictive models with training set and identify patients with SSI in the no label test set. Lab test results are vital resources that guide healthcare providers make decisions about all aspects of surgical patient management. Many machine learning models were developed after pre-processing and imputing the lab tests data and only the top performing methods are discussed. Overall, RANDOM FOREST algorithms performed better than Support Vector Machine and Logistic Regression. Using a set of 74 lab tests, with RF, there were only 4 false positives in the training set and predicted 35 out of 50 SSI patients in the test set (Accuracy 0.86, Sensitivity 0.68, and Specificity 0.91). Optimal ways to address healthcare data quality concerns and imputation methods as well as newer generalizable algorithms need to be explored further to decipher new associations and knowledge among laboratory biomarkers and SSI.

DOI: 10.4018/978-1-7998-1204-3.ch022

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Surgical Site Infections are one of the most common Hospital Acquired Infections (HSI). The United States based Centers for Disease Control and Prevention (CDC) has developed criteria, as part of its SSI surveillance program, that define surgical site infection (SSI) as any infection that occurs at or near the surgical incision after an operative procedure within 30 days of the procedure, expanded up to 90 days if prosthetic material is implanted at surgery (Horan, Gaynes, Martone, Jarvis, & Emori, 1992). SSIs are commonly confined to the incision site even though, they can extend deep into the underlying tissues. SSI is a significant burden to the post-operative patients, overall responsible for up to 30% of all hospital acquired infections and complicates 23% of abdominal surgeries (Magill et al., 2012; Aga et al., 2015; de Lissovoyet al., 2009). SSI contributes to increased morbidity, poor quality of life, mortality (up to 3%), prolonged hospital stay (up to 10 days), increased readmission rate and healthcare expenditure (Owens, Barrett, Raetzman, Maggard-Gibbons, & Steiner, 2014). There are many known and yet-to-be-known factors that play a role in SSI and a clear understanding of those factors are needed to reduce SSI and improve surgical outcomes (Lawson, Hall, & Ko, 2013; Ke et al., 2017).

LABORATORY TESTS IN THE ERA OF BIG DATA

One of the time-tested resource that has played a significant role in helping providers make decisions about choosing patients, readiness for surgery as well as intra and post-operative management are laboratory test results, such as blood counts, blood sugars including metabolic panels and inflammatory markers. These lab test results are indispensable for all aspects of patient management. The information content in lab test results could be a valuable resource to decipher unknown associations between various post-operative complications such as SSI.

Widespread adoption of Health Information Systems (HIS) including Electronic Health Records (EHR) and laboratory Information Systems has generated massive amounts of both quantitative data (e.g., laboratory values) and qualitative data (e.g., text based clinical notes), partly fueled by funding, by the Health Information Technology for Economic and Clinical Health Act of 2009. It is but inevitable that the big data management and analytics that are successfully applied in retail sales and other domains will be researched in healthcare (Murdoch & Detsky, 2013). Recent advances in computational machine learning methods and predictive modeling could be applied to analyze the information content of those lab tests and draw valuable knowledge from any associations with SSI (Sohn et al., 2017). But, clinical laboratory data (and most healthcare data) is inherent with various constraints that need to be addressed and pre-processed before applying machine learning algorithms. The sampling of lab tests is purely based on the clinical condition of the patient, as assessed by the providers, and thus the temporality of sampling is extremely variable in time/ space as well as the type of tests done. Some tests are usually done just once (e.g., ABO blood grouping) and some others are repeated multiple times (e.g., glucose, electrolytes, hemoglobin). Other than in research settings, only a core set of lab tests are done routinely on most patients during day to day clinical care and a subset of patients may be tested for disease specific lab biomarkers. This variability of sampling, both the type/ number of tests as well as temporality (time and space) contrasts with the statistical and predictive analytic tools which very much rely on systematic and regularly sampled data.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Choosing the relevant and representative features (lab tests) for a given data modeling task can be challenging, requiring significant domain expertise. Modeling with all or most of the represented laboratory tests in the patient population will have to address many missing values in the data matrix. Removing patients with missing data may lead to significant loss of sample size as well as inaccurate assessments and predictions. Thus, the clinical lab data could be sparse with many missing values in the dataset to be consumed by the predictive/statistical tools that have led to experiments with various imputation methods. Imputation methods have been successfully applied to simulate regular sampling to address the missing values. There are many imputation methods and ideal methods are still being sought and researched (Hu et al., 2017).

In this project, as part of a challenge by the American Medical Informatics Association Knowledge Discovery and Data Mining Working Group (AMIA KDDM WG), to develop better predictive models, the dataset had similar concerns. Researchers discuss the data preprocessing, imputation methods used and the various machine learning techniques – Random Forest (RF), Logistic Regression (LR) and Support Vector Machine (SVM) applied to develop predictive models.

METHODS

The data provided had all the features and analytic complexities of a big data - volume, variety, veracity, except that it was a static set of retrospective data lacking the 'velocity' of a big data. Cleaning and formatting the data for consumption by ML tools and choosing the clinically important and relevant lab tests as features for predictive modeling of the already dimension reduced data was the challenge during the pre-processing phase. Dimensions that are commonly associated with healthcare data such as presenting complaints, symptoms and signs, diagnoses and other investigations such as imaging, procedures, medications and clinical outcomes were not part of the data provided. Researchers propose a semi-automated computational method of the model development. Figure 1 illustrates the overall process flow of the proposed model. The training data set provided included a vast number of investigation results spanned over many days both before and after the day of surgery. Further, for some of the relevant investigations, data was sparse across the patient population and there were many missing values. Thus, the first part of the proposed method was to preprocess the data based on investigations, time frame and imputing the missing values. In the next step relevant features were extracted and a training model was built using appropriate and common machine learning algorithms. The performances of selected machine learning algorithm based models were evaluated to classify patients with SSI and patients without. The steps were repeated with the unlabeled test data, and after preprocessing the developed training models were applied to test data to discriminate and develop a list of potential SSI patients.

DATA DESCRIPTION

The training dataset had lab tests done on 909 patients with 183 patients having SSI and 726 not having SSI, and the test set had 228 patients with 50 having SSI, although data was blinded without SSI labels (Table 1).

As shown in Table 1, in the training set, males (54% of the population) were more prone to SSI compared to females (23.34% vs 16.22%). The age distribution was from 5 years to 110 years and most of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Preprocessing steps for feature selection

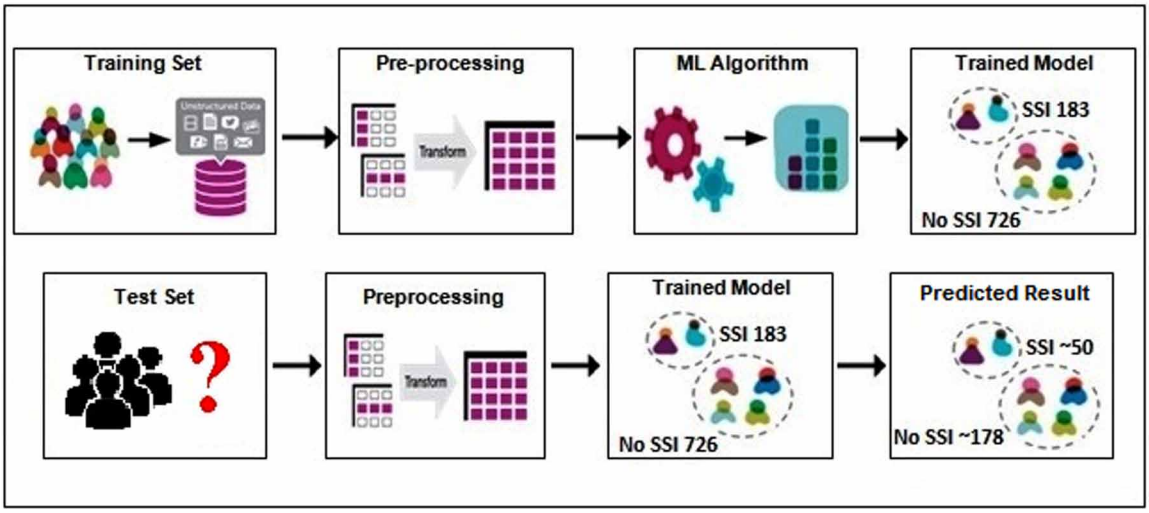


Table 1. Demographic distribution of patients in the training and test datasets

File Types	The Number of Cases Infected	The Number of Controls	Males	Females	Males Infected	Females Infected	Total Number
Training	183 (25.5%)	726	490 (53.91%)	419 (46.09%)	115 (23.34%)	68 (16.22%)	909
Testing	50 (28.08%)	178	123 (53.94%)	105 (46.05%)	?	?	228
	Who are those?						

the patients who underwent surgery were in the age group 50 to 80 years with the corresponding increase in infection rate. A small number of patients below age 19 and above 94 didn't have SSI (Figure 2).

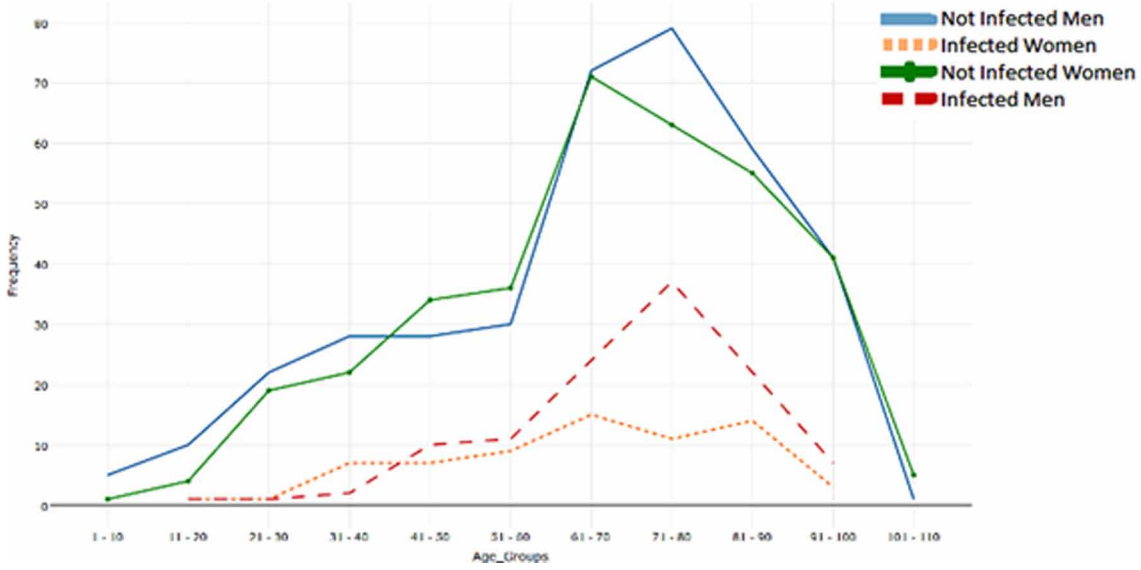
With regards to the timing of blood tests to surgery, most patients had tests done before, on the day of surgery and in the post-operative period. About 32 in the training set (2 with SSI) and 6 patients in the test set had tests done only in the post-operative period, whereas 18 in the training set and 1 in the test set had tests done only during the pre-operative phase.

The data was from the department of Gastrointestinal Surgery at University Hospital of North Norway. The lab test types were in Norwegian and there was no associated metadata. There was a total of 504,675 blood test result records in the dataset having 811 unique lab test types. Not all the 811 tests were done on all the patients, probably based on clinical indications, for example, a tumor marker such as Carcinoembryonic Antigen (CEA) was done on patients with neoplasms. Only a small core subset (~10%) of them conducted routinely on significant patient population, at variable intervals, often repeated many times, and associated with many results.

Patients with lab test result values that were outside the normal range (abnormal) in the dataset were flagged with a star ("*"). However, the marking of test result values was variable with some high and low results not marked as abnormal (e.g., CPR of 295 not marked as abnormal with a star, normal <5)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Age and gender distribution of patients with SSI and no SSI



compared to what is available in the literature, even though normal ranges were not provided by the organizers. The data spans over 15 years from 1994 till 2013 and over time, and it represents many methods to measure (e.g., biochemical assay, mass spectrometry) and different specimen sources (e.g., venous, arterial and capillary blood) for some of the molecules, with corresponding differing normal ranges. There were 5 different types of CRP (CRP, CRP (H, N), CRP sensitive, CRP TMS, CRP semikvantitativt). Most of them were in small numbers and were excluded during preprocessing, but, CRP (H, N) was in significant number, was treated as a separate test from CRP and in fact was ranked within the top ten relevant tests by all the three classifiers (LR, RF and SVM). Some of the data elements were of varied data types (e.g., date), adding to the complexity of handling such data. After cleaning the data, the major task was feature selection, to eliminate features that were identified clinically and statistically irrelevant and imputation to address the missing values to resemble regularly sampled data to be consumable by statistical/ predictive modeling tools to boost their performance.

DATA PRE-PROCESSING

Feature Selection

A series of procedures were identified and followed for feature selection. First, all the abnormal result values were extracted by pattern matching the star flagging (“*”) along with their test types. This gave a list of test types for which abnormal values were observed. If a test was not out of range during the study period (15 years) for many patients, it is possible that its potential to discriminate normal from abnormal is limited and may not add value with classifying patients as having an infection or no infection. This list of test types was reviewed and compared to the original test types by a clinician (PS) and observed that all the clinically important and relevant test types were in the list. Only these test types

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

were used for further data processing and rest of the test types were eliminated as they were deemed clinically irrelevant. Using the list of processed lab test types, all Patient IDs (PIDs) whose lab test types matched were identified to create a subset of PIDs. All the patients with SSI were further identified by the patient labels having 1 or 2 in the 'infection' column in the original dataset.

In the next step, after formatting the date data type, the dataset had records of a patient whose tests spanned 15 years before and after the surgery. Results of tests done more than 1 year before the day of surgery are unlikely to help with classifying infected and not infected, and lab test type value records of one year before and after the date of surgery (-1 to +1 year from surgery) were retained. The number of test types reduced to 174. Further if only one or very few patients had a test type, then that test type is probably not relevant for all. If a test has been performed on at least 10 patients with SSI, then such test types were preserved, and the remaining test types were eliminated which reduced the number of test types (features) to 107. This was further validated by a clinician. Researchers experimented with using mode or standard deviation for further feature selection (instead of 10 patients with SSI), but, no two test result values were same and selecting features based on two standard deviations resulted in the loss of clinically significant test types such as hemoglobin. The 107 test types were used as the training dataset to model using RF, SVM and LR and other records were eliminated. In a second model, the time duration was further reduced to 30 days before and after surgery, excluding the day of surgery, and the number of features further reduced to 74 test types, yet retaining all clinically relevant test types.

An analysis of population mean results of some of the frequently done tests distinctly demonstrated discriminatory potential, between patients with SSI and no-SSI, especially during the post-operative period (Figure 3). These mean values contributed in training the model taking a cumulative value for each test done on a patient for the considered span of time.

A data matrix was developed (Table 2), with test types in the x-axis and patient (PID) represented in the y-axis. For each PID, mean of values of Test (Test 1…Test m) for the patient for each test type is plotted. Not Available (NA) was substituted for records where test type was not conducted for a patient. For example, if HbA1c test is not done on PID 1234 then the table has a value NA.

Imputation Methods

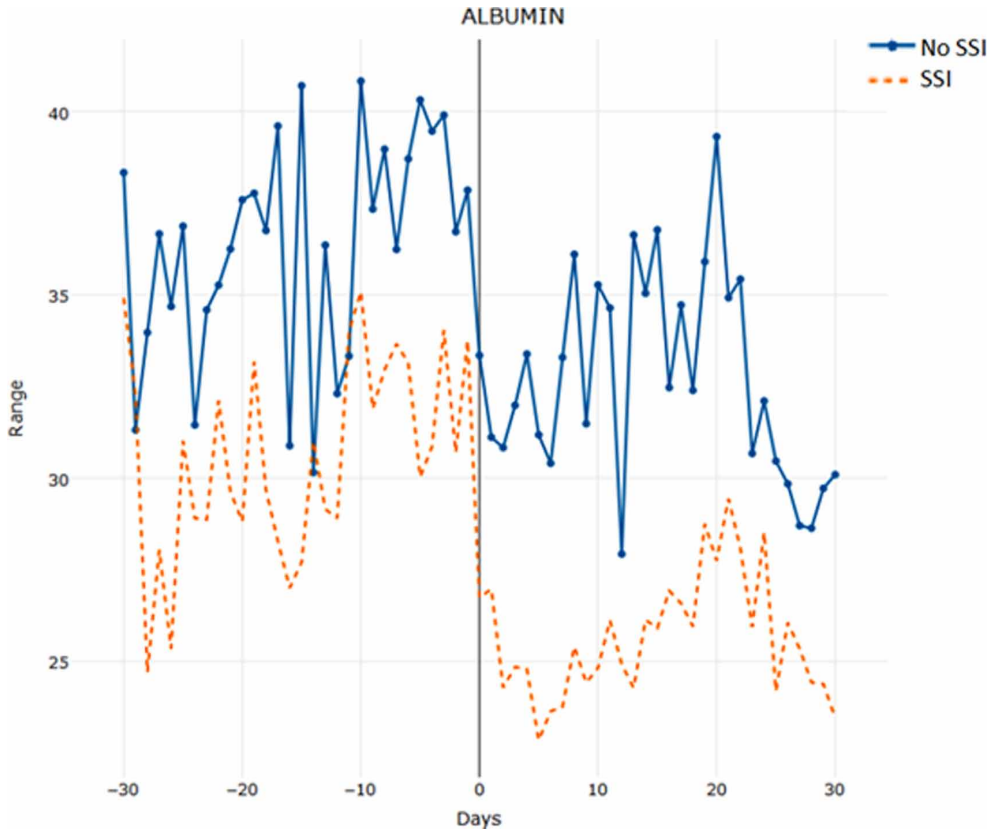
To address the missing test type results, the within range (normal) values from training dataset was determined by extracting all test result values not being flagged with a star ("*"). The test types for each of these values were extracted and the mean of normal values for each test type was computed. This

Table 2. Data Matrix after preprocessing for ML consumption

PID	Test 1	Test 2	Test 3	...	Test M	Infected
						(class label)
1	23	NA	NA	...	90	1
2	NA	56	78	...	NA	0
3	45	67	NA	...	NA	1
.	NA	NA	73	...	98	1
.	20	50	NA	...	NA	0
N	NA	54	NA	...	80	0

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Sample Mean for Albumin 30 Days Pre-Post Surgery



mean was imputed for all missing lab test type result values (NA), both with SSI and without, in the final training dataset. The mean of 'normal' test result values were used to impute all the patients as the 'abnormal' test result values had a wider range, including a few outliers with extreme values. Soguero-Ruiz et al. (2015) proposed a way to impute values for the missing data but they did not consider the abnormal cases while imputing.

Machine Learning Methods

Machine learning is a branch of artificial intelligence and ML algorithms iteratively tries to understand and learn the pattern space and its associated labels. Once the pattern space is understood, they predict the incoming test patterns and assign appropriate labels. Supervised machine learning is tasked with inferring a function from the labeled training dataset. A detailed description of SVM, Random Forest and Logistic Regression is beyond the scope of this manuscript and a brief review is provided below.

SVM is a supervised machine learning technique. Training examples are mapped to points in high dimensional space using the values of the features. SVM computes a hyperplane and separates the positive points from negative points with the maximum margin. Test examples are classified as positive or negative based on which side of the hyperplane they belong. Random Forest is an ML method that creates an ensemble of decision trees, using dataset bagging, feature set bagging and other techniques.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bagging is a procedure to reduce the statistical variance in the training data during tree splitting. RF uses averaging (voting) of the constituent decision tree classifiers to improve the accuracy of predictions and control over-fitting (Breiman, 2001). Logistic Regression is a widely used generalized linear classification model.

From the literature (Sanger et al., 2016; Van Esbroeck, Rubinfeld, Hall, & Syed, 2014; Gbegnon, Monestina, & Cromwell, 2014) it's evident that machine learning approaches are better suited for predicting surgical site infections. Random Forest (Liaw & Wiener, 2002; Statnikov, Wang, & Aliferis, 2008; Kuhn, 2012), SVM (Bennett & Campbell, 2000; Chang& Lin, 2011; Cortes & Vapnik, 1995; Schölkopf, Smola, Williamson, & Bartlett, 2000) and Logistic Regression were used to train, test and compare the accuracy of prediction by clinically relevant features.

Many models were developed, and researchers discuss some of the best performing models. Random Forest was the preferred method for various reasons:

- RF is an ensemble learning method for classification, regression and other tasks and randomization at many (data and model) levels.
- The data provided was already split 4:1 into training and test sets by the organizers and further splitting data into training/ test sets is usually not required for RF.
- Early modeling with the training data, RF performance was better than SVM and LR and researchers preferred RF for most of the modeling work.

RESULTS

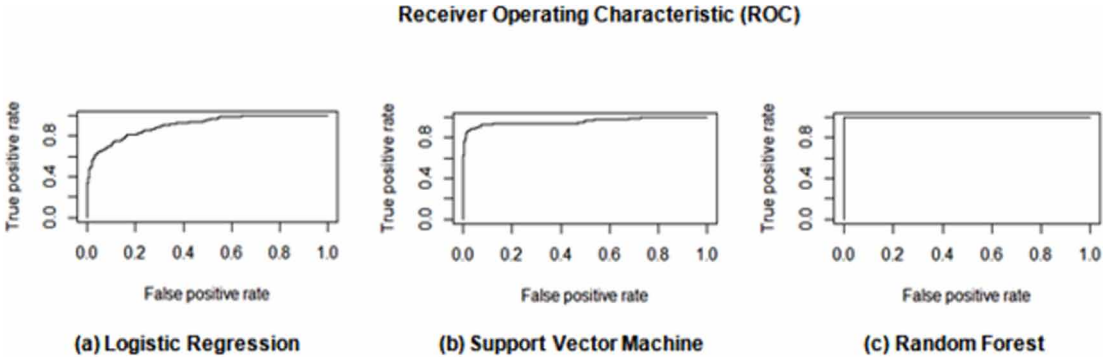
After preprocessing and imputing the data as mentioned above, 1 year pre and post-surgery data with 107 test types were used to train models using three different ML methods SVM, LR and RF. Predictive models were evaluated with respect to two criteria namely classifier accuracy and clinical relevance of the lab test types (features) indexed by the algorithms.

The Receiver Operating Characteristic (ROC) curves are one of the most commonly used and key statistical tool to evaluate the accuracy of diagnostic tests in healthcare. They are often used to establish relationships between a continuous variable for a binary outcome. It represents 1-specificity and sensitivity pair, plotted as a graph for all possible threshold values, demonstrating the trade-off between false positive rate (1-specificity) and true positive rate (sensitivity) for the different possible cutpoints of a test (Yin & Vogel, 2017). The closer the curve follows the left-hand and top borders of the ROC space, the more accurate the test and the area under the curve (AUC) is a measure of test accuracy. Figure 4 shows the ROC of training data from all the three classifiers. It is evident from the ROC plot that false positive rate for logistic regression is high compared to SVM and RF.

Even though accuracy was variable between classifiers (SVM = 91.32%, LR = 89.24%, RF = 99.89%), some of the important lab test types selected are clinically relevant and already been implicated in the literature as predictors of SSI. The CRP, including CRP (H, N), Albumin and Platelets are identified by all the three methods as of high relevance (Ge et al., 2017). RF produced the best results with the training set with an accuracy of 99.89%, Sensitivity (SN) of 0.9945, Specificity (SP) of 1.0 and 0 False Positives. RF also predicted the most clinically relevant lab tests as evaluated by a clinician when compared to SVM and LR classifiers. Table 3 shows the comparison of different ML models with top relevance indexed features and their clinical importance among the infected and non-infected cohorts of patients.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Receiver operating characteristics (ROC) for the training dataset using different classifiers



In the model using RF and top 10 features on test dataset, results of which were submitted to the organizers for validation, about 55 patients were predicted as having SSI, 5 more than anticipated. Once the test set SSI labels were made available, only 33 out of 55 patients were found to have SSI, with prediction Accuracy of 82.89%, SN 0.66, SP 0.87, Positive Predictive Value (PPV) 0.6 and Negative Predictive Value (NPV) 0.901. Among the many models the researchers developed, the 30 days pre-post, excluding the day of surgery (homeostatic perturbance are high on the day of surgery), produced the best results with 74 lab tests, correctly predicting 35 patients as having SSI in the test set (Accuracy 86.4%, SN 0.7, SP 0.91, PPV 0.686, NPV 0.915). The confusion matrix comparing the output of the above models, using RF classifier on test dataset, with true SSl labels are shown in Table 4 and test dataset ROC curves are shown in Figure 5. Even though questions about overfitting were raised, especially with RF, it was deemed unlikely with a training sample size of 897 (+/- 1 year) and 823 (+/- 30 days) and consistent performance by modeling with RF using various combinations of features (test types).

To check the least number of lab tests that may be able to predict SSI, researchers also developed models based on various combinations of the top 5 lab tests (CRP, Albumin, Leukocytter, Haemoglobin and Trombocytter), but, none of the combinations improved the predictive performance. In literature, Prealbumin and Procalcitonin have been found to be useful for predicting SSI. In the provided dataset, there were only 86 patients who had Prealbumin tests, 18 of them had SSI and the sample mean did not show any significant variation between patients with and without SSI. There were no Procalcitonin test results in the dataset (De Magistris et al., 2016).

CONCLUSION

The methods researchers followed for feature selection and imputation of missing data are implemented for the first time considering all the longitudinal lab test results of the patients. The pre-and-post-operative phases were not considered separately as the goal was to predict patients with SSI in the test dataset, fully utilizing the variations in test type results between patients, irrespective of the phase of surgery. This is particularly of importance as most patients were already ill preoperatively, requiring surgery, with abnormal blood test results (e.g., Hemoglobin, CRP, Albumin and Phosphate). Other contributing factors for abnormal test results such as comorbidities, blood and blood product transfusions, medications and other patient dimensions were not available. The performance of RF classifier was the best on the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Comparison of different classifiers with their ranked feature's clinical importance

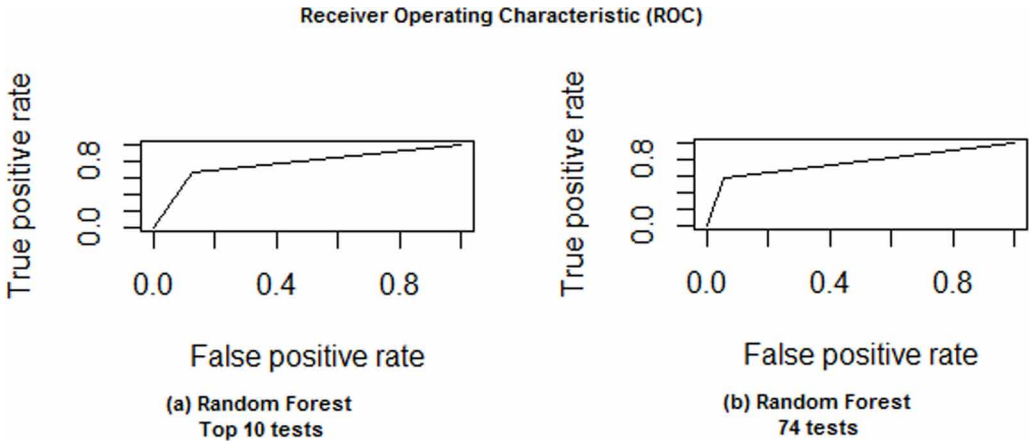
SVM				Random Forest				Logistic Regression			
Test Types	**Importance %**	**Not-infected**	**Infected**	**Test Types**	**Importance %**	**Not-infected**	**Infected**	**Test Types**	**Importance %**	**Not-infected**	**Infected**
CRP	100	680	172	CRP	100	680	172	CRP	100	680	172
Leukocytter	79.36496823	718	183	Albumin	96.56973526	648	174	CRP (H,N)	85.08596132	160	55
Basofile	75.94913445	319	106	Hemoglobin	62.55242131	723	183	Antitrombin III	73.49365435	74	44
Basofile%	73.18005986	303	107	Trombocytter	59.97843623	644	178	Albumin	70.13943734	648	174
CRP (H,N)	67.72505295	160	55	Leukocytter	45.1093067	718	183	CK	68.0020328	464	114
SR	67.02228212	485	114	CRP (H,N)	32.442271	160	55	Trombocytter	60.65919433	644	178
Trombocytter	66.94439096	644	178	Natrium	31.75899069	721	183	Cystatin C	56.92875853	56	30
D-dimer	66.10684217	192	66	PT-INR	30.87164514	454	139	IgA Kvantitering	54.78700874	215	24
PT-INR	65.94493357	454	139	Kalium	30.29122556	709	181	Kalium	51.59609714	709	181
GT	63.36139749	652	174	Calsium	30.21409708	520	140	Fritt-T4	45.73491293	477	106

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Confusion matrix comparing output of two models using RF classifier on test dataset with true SSl labels (R Statistical Package)

Predictive Performance Measures	1 Year Pre-Post (Top 10 Features) vs AMIA Test Dataset SSI Labels			30 Days Pre-Post (74 Features) vs AMIA Test Dataset SSI Labels		
		True SSI Labels			True SSI Labels	
	Model 1	1 (SSI)	0 (No-SSI)	Model 2	1 (SSI)	0 (No-SSI)
	(+/- 1 year)			(+/- 30 Days)		
	1 (SSI)	33	22	1 (SSI)	35	16
	0 (No-SSI)	17	156	0 (No-SSI)	15	162
Accuracy	0.828			0.864		
Sensitivity	0.66			0.7		
Specificity	0.876			0.91		
Positive Predictive Value	0.6			0.686		
Negative Predictive Value	0.901			0.915		

Figure 5. Receiver operating characteristics (ROC) of the results of two models using RF classifier on test dataset with true SSl labels



processed data compared to SVM and LR, and identified the most clinically relevant test types (features) such as CRP, albumin, hemoglobin, platelets and leukocytes. As anticipated, the performance of models built on combined pre and post-operative data was better than preoperative data alone (1 year). This suggests that the information content and predictive value of certain lab tests and their longitudinal values are more robust at classifying patients with and without SSI, with continued monitoring during both pre and post-operative phases. Previous studies with machine learning tools have shown that, in general, thrombocytes, ALP, CRP, albumin, creatinine and leukocytes, were most useful features in predicting SSI (Soguero-Ruiz et al., 2015). The implicated lab tests also validate the many years of established clinical practice of dependence on those biomarkers to monitor SSI.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Feature selection is critical in any predictive modeling, and utilizing the top 10 features selected from the training set in the first model (1-year pre-post), RF classified the test set identifying SSI patients (55) more towards the target infected population of 50. However, the first model identified only 33 true positives in the test set (Accuracy 82.89%, SN 0.66, SP 0.87, PPV 0.6, and NPV 0.9). In the best performing model (30 days pre-post), 74 features were used, predicted about 51 patients as having SSI in the test set, among which 35 were true positives (Accuracy 86.4%, SN 0.7, SP 0.91, PPV 0.68, NPV 0.91). This suggests that given the test results, 91% of the time you can predict who may not get SSI whereas 68% of the time those who may get SSI. As anticipated, there was a reduction in accuracy, sensitivity, specificity, PPV and NPV between the training set and test set. However, the reduction in sensitivity and PPV are more evident compared to specificity and NPV. The decay in performance between training and test sets is probably related to many factors. The same feature selection methods used for training set (considering out of range lab tests only, 1 year and 30 days post-surgery, and tests done on 10 infected patients and above) could not be applied to test set as it would have resulted in eliminating patients in the test set. Excluding patients in the test set was not an option as researchers did not want to reduce the test set sample size and there also existed a potential of eliminating true positive SSI patients from the test set. Small test sample size, and imputation of feature values with the mean of the normal range of test types, could be other reasons. Researchers plan to test further models in RF as well as SVM with larger sample size to address many questions including an optimal set of features and imputation methods that improve predictive accuracy. Advances in neural networks may afford more generalizable algorithms to model and improve predictive performance.

It is clear from various experiments that the lab tests are important predictors of SSI that could help surgeons predict who may not or may get SSI, so that they can make informed decisions about appropriate preventative measures. With big data infrastructure and analytic tools, the above methods could be used to develop local live dynamic models, to predict at the bedside, who may not and who may get SSI, using lab test types and other patient parameters (Hu et al., 2015). This could be a valuable resource at a time where hospital acquired infections, including SSI, are a global menace further fueling the development of antibiotic resistance.

ACKNOWLEDGMENT

We would like to thank the AMIA Knowledge Discovery and Data Mining (AMIA - KDDM) Work Group Data Competition Program Committee, 2016, for organizing and conducting the competition and the University Hospital of North Norway for providing the data creating an opportunity for us to participate in the project.

REFERENCES

Aga, E., Keinan-Boker, L., Eithan, A., Mais, T., Rabinovich, A., & Nassar, F. (2015). Surgical site infections after abdominal surgery: Incidence and risk factors. A prospective cohort study. *Infectious Diseases*, *47*(11), 761–767. doi:10.3109/23744235.2015.1055587 PMID:26114986

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bennett, K. P., & Campbell, C. (2000). Support vector machines: Hype or hallelujah? *ACM SIGKDD Explorations Newsletter*, *2*(2), 1–13. doi:10.1145/380995.380999

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi:10.1023/A:1010933404324

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 27. doi:10.1145/1961189.1961199

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018

de Lissovoy, G., Fraeman, K., Hutchins, V., Murphy, D., Song, D., & Vaughn, B. B. (2009). Surgical site infection: Incidence and impact on hospital utilization and treatment costs. *American Journal of Infection Control*, *37*(5), 387–397. doi:10.1016/j.ajic.2008.12.010 PMID:19398246

De Magistris, L., Paquette, B., Orry, D., Facy, O., Di Giacomo, G., Rat, P., ... Ortega-Deballon, P. (2016). Preoperative inflammation increases the risk of infection after elective colorectal surgery: Results from a prospective cohort. *International Journal of Colorectal Disease*, *31*(9), 1611–1617. doi:10.100700384-016-2620-8 PMID:27357511

Gbegnon, A., Monestina, J., & Cromwell, J. W. (2014). Machine Learning Algorithm for Accurate, Automated, Real-Time Prediction of Surgical Site Infections using EHR Data. *The Journal of Surgical Research*, *186*(2), 527. doi:10.1016/j.jss.2013.11.380 PMID:24559584

Ge, X., Cao, Y., Wang, H., Ding, C., Tian, H., Zhang, X., ... Li, N. (2017). Diagnostic accuracy of the postoperative ratio of C-reactive protein to albumin for complications after colorectal surgery. *World Journal of Surgical Oncology*, *15*(1), 15. doi:10.118612957-016-1092-1 PMID:28069031

Horan, T. C., Gaynes, R. P., Martone, W. J., Jarvis, W. R., & Emori, T. G. (1992). CDC definitions of nosocomial surgical site infections, 1992: A modification of CDC definitions of surgical wound infections. *Infection Control and Hospital Epidemiology*, *13*(10), 606–608. doi:10.1017/S0195941700015241 PMID:1334988

Hu, Z., Melton, G. B., Arsoniadis, E. G., Wang, Y., Kwaan, M. R., & Simon, G. J. (2017). Strategies for handling missing clinical data for automated surgical site infection detection from the electronic health record. *Journal of Biomedical Informatics*, *68*, 112–120. doi:10.1016/j.jbi.2017.03.009 PMID:28323112

Hu, Z., Simon, G. J., Arsoniadis, E. G., Wang, Y., Kwaan, M. R., & Melton, G. B. (2015). Automated Detection of Postoperative Surgical Site Infections Using Supervised Methods with Electronic Health Record Data. *Studies in Health Technology and Informatics*, *216*, 706–710. PMID:26262143

Ke, C., Jin, Y., Evans, H., Lober, B., Qian, X., Liu, J., & Huang, S. (2017). Prognostics of surgical site infections using dynamic health data. *Journal of Biomedical Informatics*, *65*, 22–33. doi:10.1016/j.jbi.2016.10.021 PMID:27825798

Kuhn, M. (2012). Contributions from Jed Wing SW, Andre Williams, Chris Keefer and Allan Engelhardt. caret: Classification and Regression Training (R package version 5.15-023).

Lawson, E. H., Hall, B. L., & Ko, C. Y. (2013). Risk factors for superficial vs deep/organ-space surgical site infections: Implications for quality improvement initiatives. *JAMA Surgery*, *148*(9), 849–858. doi:10.1001/jamasurg.2013.2925 PMID:23864108

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Magill, S. S., Hellinger, W., Cohen, J., Kay, R., Bailey, C., Boland, B., ... Goraczewski, L. (2012). Prevalence of healthcare-associated infections in acute care hospitals in Jacksonville, Florida. *Infection Control and Hospital Epidemiology*, *33*(03), 283–291. doi:10.1086/664048 PMID:22314066

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Journal of the American Medical Association*, *309*(13), 1351–1352. doi:10.1001/jama.2013.393 PMID:23549579

Owens, P. L., Barrett, M. L., Raetzman, S., Maggard-Gibbons, M., & Steiner, C. A. (2014). Surgical site infections following ambulatory surgery procedures. *Journal of the American Medical Association*, *311*(7), 709–716. doi:10.1001/jama.2014.4 PMID:24549551

R Core Team. (2013). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from http://www.R-project.org/

Sanger, P. C., van Ramshorst, G. H., Mercan, E., Huang, S., Hartzler, A., Armstrong, C. A., ... Evans, H. L. (2016). A Prognostic Model of Surgical Site Infection Using Daily Clinical Wound Assessment. *Journal of the American College of Surgeons*, *223*(2), 259–270.e2. doi:10.1016/j.jamcollsurg.2016.04.046 PMID:27188832

Schölkopf, B., Smola, A. J., Williamson, R. C., & Bartlett, P. L. (2000). New support vector algorithms. *Neural Computation*, *12*(5), 1207–1245. doi:10.1162/089976600300015565 PMID:10905814

Soguero-Ruiz, C., Fei, W. M., Jenssen, R., Augestad, K. M., Álvarez, J. L. R., Jiménez, I. M., ... Skrøvseth, S. O. (2015). Data-driven Temporal Prediction of Surgical Site Infection. *AMIA ... Annual Symposium Proceedings - AMIA Symposium. AMIA Symposium*, 1164. PMID:26958256

Sohn, S., Larson, D. W., Habermann, E. B., Naessens, J. M., Alabbad, J. Y., & Liu, H. F. (2017). Detection of clinically important colorectal surgical site infection using Bayesian network. *The Journal of Surgical Research*, *209*, 168–173. doi:10.1016/j.jss.2016.09.058 PMID:28032554

Statnikov, A., Wang, L., & Aliferis, C. F. (2008). A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinformatics*, *9*(1), 1. doi:10.1186/1471-2105-9-319 PMID:18647401

Van Esbroeck, A., Rubinfeld, I., Hall, B., & Syed, Z. (2014). Quantifying surgical complexity with machine learning: Looking beyond patient factors to improve surgical models. *Surgery*, *156*(5), 1097–1105. doi:10.1016/j.surg.2014.04.034 PMID:25108343

Yin, J., & Vogel, R. L. (2017). Using the ROC Curve to Measure Association and Evaluate Prediction Accuracy for a Binary Outcome. *Biometrics & Biostatistics International Journal*, *5*(3), 1. doi:10.15406/bbij.2017.05.00134

This research was previously published in the International Journal of Big Data and Analytics in Healthcare (IJBDAH), 3(1); edited by Mu-Yen Chen; pages 13-26; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 23
Towards Semantic Interoperability in Health Data Management Facilitating Process Mining

Barbara Traxler
University of Applied Sciences Upper Austria, Austria

Oliver Krauss
University of Applied Sciences Upper Austria, Austria

Emmanuel Helm
University of Applied Sciences Upper Austria, Austria

Andreas Schuler
University of Applied Sciences Upper Austria, Austria

Josef Kueng
Johannes Kepler University Linz, Austria

ABSTRACT

As an evidence-based business process analysis method, process mining can be used to investigate variations in delivery of care. Existing approaches are only based on one data source. A variety of data sources means different domain languages and understanding, special processes workflows in various organizations, varying documentation with different goals and different designations and varying use of coding systems. This article describes a modular, rule-based information extraction algorithm based on CDA and compares it to a proprietary healthcare reference model approach and a resource-based extraction of healthcare data using the new standard FHIR. All three approaches can be used to derive models to extract clinical and patient pathways. Similarities and differences according to interoperability and process mining tasks are described. It is concluded that standards-based approaches allow for more interoperability and can be used for a wide range of systems to provide process insight, thus facilitating better healthcare management across institutional boundaries.

DOI: 10.4018/978-1-7998-1204-3.ch023

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Efficient and patient-centered delivery of care across different hospital settings and healthcare providers requires intensive coordination between concerned organizations. Clinical pathways are used as management tools to define the best process in a healthcare organization, using the best procedures and timing, to treat patients with specific diagnoses or conditions according to evidence-based medicine (Panella, Marchisio and Di Stanislao, 2003). (Partington et al., 2015) propose the application of process mining as an evidence-based business process analysis method to gain insight into and enable comparative analysis of clinical practice and delivery of care across different organizations.

The objective of this publication is to give an overview of available models that facilitate process mining in Healthcare. Since information on (proprietary) models is often lacking, or not publicly available, the scope of this research is limited to models based on international standards, which are free of charge.

Existing approaches which use process mining for healthcare management are only based on data sources within one organization. More precisely the formal representation and the semantics (including code systems and value sets) of the different data sources are basically the same (Partington et al., 2015; Mans et al., 2008). To investigate variations in clinical practice and delivery of care across different healthcare domains, various data sources have to be taken into account (as shown in Table 1): Integrating the Healthcare Enterprise (IHE) and Health Level Seven (HL7) Clinical Document Architecture (CDA) are used in wide parts of the US and Europe as a technically interoperable solution to exchange healthcare data between healthcare providers (Franz et al. 2011). The use of standardized mechanisms increase the interoperability of information systems and minimize integration efforts (Norgall, 2003, Sunyaev et al., 2008). Nevertheless, there is still a lot of healthcare data exchange done using proprietary systems. New standards like HL7 Fast Healthcare Interoperability Resources (FHIR) try to cover this gap and allow a resource efficient data transfer (publication). OpenEHR and the use of ontology-based archetypes is another approach to allow interoperable healthcare data management. Since it uses SNOMED-CT (IHTSDO) to support automated clinical processes, which is subject to a fee, it is not taken into account in this article. (Mans, van der Aalst and Vanwersch, 2015) describe a (proprietary) healthcare reference model (HRM) that aims to help locating the needed data in healthcare information systems and thus facilitate the data extraction for process mining.

A variety of data sources means different domain languages and understanding, special processes and workflows in various organizations, varying documentation with different goals, varying structures in documentation, different designations and varying use of standards and coding systems. Thus, to extract data out of the (distributed) data sources and to transform them into the structure needed for process mining algorithms, preprocessing steps are necessary. This work describes and evaluates different approaches as a preceding step to enable the development of clinical pathways and medical guidelines.

Preceding Steps for Process Mining

Process mining algorithms work on event logs with a certain structure. Event logs must contain only data related to a single process and it must be ensured that all events in the log can be related to this process. Moreover, each event in the log must represent an activity and refer to a single process instance (case). The Extract, Transform and Load (ETL) steps preceding the actual process mining tasks describe: (a) extraction data from outside sources, (b) transforming it to fit operational needs (dealing with syntactical

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Overview of existing approaches that facilitate process mining, and reasons for in/exclusion in this publication

Model	Included / Excluded	Reason
FHIR	yes	Defining Processes and Documenting their Execution are part of the standard
HRM	yes	Used in scientific literature to enable process mining. Proprietary HIS models are often transposed into HRM for analysis.
openEHR	no	Use of SNOMED-CT to support modeling of automated clinical processes
Proprietary approaches	no	Publicly available information not sufficient
RIM (CDA/IHE)	yes	HL7 RIM is widely used in healthcare, including IHE profiles that define process(steps) for medical use-cases.

and semantical issues while ensuring predefined quality levels), and (c) loading it into the target system, e.g. a data warehouse or relational database (Van der Aalst, 2011). Thus, interoperability between various healthcare data sources is a prerequisite to allow for significant process mining.

BACKGROUND

Interoperability Between Healthcare Providers

Integrated care and the achievement of high quality healthcare over institutional borders require three levels of interoperability between different healthcare providers (Heitmann and Gobrecht, 2009, Serrano et al., 2015):

- **Technical Interoperability:** Exchange of meaningful, actionable information between two or more systems across organizational boundaries
- **Semantic Interoperability:** A shared understanding of the exchanged information
- **Process Interoperability:** An agreed expectation for the response to the information exchange and requisite quality (reliability, fidelity, security) of services and processes.

Since technical interoperability has been the focus of standards organizations, alliances and consortia for many years, standards and implementations supporting this level of interoperability are generally available (e.g. IHE, HL7). Strategies for informational interoperability, however, which includes the whole area of semantic and process interoperability, are less mature (Serrano et al., 2015). Since all three levels of interoperability are necessary to achieve high quality healthcare over institutional borders, the focus of this article lies on these types of interoperability for the comparison of the various modeling approaches. Using pre-defined use cases and based on the interoperability levels described above, the HRM-, the FHIR- and the CDA-approach are analyzed how they add semantic information to the models and whether they support the ETL process.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Healthcare Reference Model

(Mans, van der Aalst and Vanwersch, 2015) use the healthcare reference model (HRM) to compare processes of two different hospitals. The model was specifically designed for the healthcare domain to locate the data in the proprietary healthcare information systems, thus facilitate the data extraction for process mining. The HRM was developed in reverse engineering the data model of a running i.s.h.med[1] hospital information system (HIS) based on the actual database table structure, expert interviews and hands-on inspection validated via interviews with HIS professionals of other hospitals (Mans, van der Aalst and Vanwersch 2015).

The HRM data model comprises 122 UML classes, which are grouped in several sub-models such as General Patient and Case Data, Radiology and Document Data and include the key elements needed for process mining (Mans, van der Aalst and Vanwersch, 2015). The developers of the HRM do not claim completeness of the model since HISs of different vendors may contain data not present in the model.

Reference Information Model

HL7 Clinical Document Architecture (CDA) is a well-established document-based exchange standard, specifically developed for the health sector to allow a uniform understanding of objects and processes in the healthcare environment. It defines the exchange of messages, document-based communications as well as cooperating services, their implementation and necessary infrastructural services (Benson and Grieves, 2016). The core is described in Refined Message Information Models (R-MIM), which are derived from the HL7 Reference Information Model (RIM), which is a generic healthcare specific information model (Kühne, 2006). The base of this model are four core classes (Act, Entity, Role and Participation) and two additional classes (ActRelationship and RoleLink). The use of RIM provides specifications to structure, type, content as well as semantics, used vocabulary and underlying processes necessary for data transfer and interoperability, following a top-down approach.

For the representation of clinical and patient pathways HL7 proposes the care plan model, for which a three-layered modelling approach was applied to allow separation of business, information and interoperability requirements (HL7, 2016):

- **Layer 1:** The Care Plan Conceptual Model identifies necessary concepts and the relationships between them. It consists of an abstract concept Plan that is associated with further concepts, e.g. Care Giver, Patient, Provider, Activity. These are directly derived from high level business requirements and thus build the foundation for the consecutive layers.
- **Layer 2:** The Care Plan Organizing Framework builds the meta-model for coordination of care interaction and collaboration and defines relationships between a subset of the concepts defined in layer 1.
- **Layer 3:** The Care Plan Logical Information Model adds data properties to the concepts defined in the preceding layers, which captures information relevant for dynamic coordination of care interactions and point in time data exchange.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The resulting care plan structure is thus applicable in a wide range of scenarios and use cases (HL7, 2016), consisting of discipline- or treatment-specific plans as well as comprehensive multidisciplinary plans, e.g. in case of tumor board review meeting. An implementation of the Care Plan Model defined by HL7 is currently developed as part of HL7s standard Fast Healthcare Interoperability Resources and the CarePlan resource respectively.

Fast Healthcare Interoperability Resources

Fast Healthcare Interoperability Resources (FHIR) is a resource-based data exchange standard for healthcare information. FHIR resources can be viewed in different formats, such as a Structural View, UML, XML, JSON and Turtle for understanding and rapid prototyping. Further, FHIR Resources contain mapping information to the HL7 RIM, defining which fields of a resource correspond to which RIM concept. All resources, which are currently defined in HL7 FHIR, are listed in the current STU 3 version (HL7 2018). FHIR follows the Pareto Principle, better known as the 80/20 rule, and thus defines only the data exchange information 80% of identified use cases require. Specializations for the other 20% that may be required, can be modeled by implementers using Extensions and Profiles (Benson and Grieves 2016). FHIR is being updated regularly, with the next release planned in spring 2018.

FHIR is organized in different levels building upon each other:

- **Level 1:** The basis of the standard, such as the API description, Data Types and Data Formats.
- **Level 2:** Security and Implementer information as well as Terminology Bindings and Conformance Resources such as Structure Definitions describing FHIR Resources.
- **Level 3:** Contains the basis for real world use-cases such as the Patient and Practitioners.
- **Level 4:** Deals with data exchange in healthcare including clinical, diagnostic, medical, financial and workflow data.
- **Level 5:** Describes clinical reasoning and contains resources enabling automation in that sector.

MODELING CONTEXT-SPECIFIC PATHWAYS

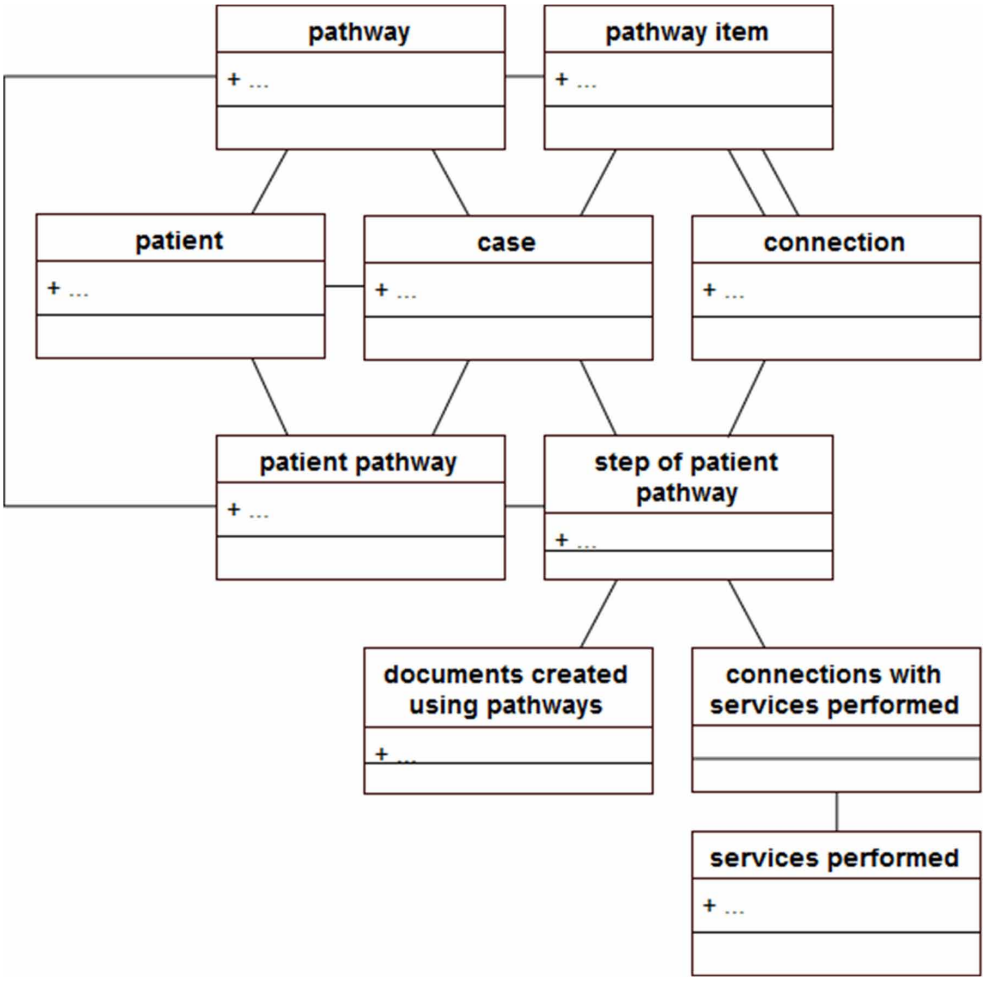
This section describes how the HRM, CDA and FHIR can be used to model clinical and patient pathways, while taking context into consideration.

Using HRM to Model Pathways

For the HRM the definition and execution of pathways is described using 10 UML classes (see Figure 1). A pathway (pathway) consists of multiple items (pathway item) which may be connected to each other (connection). A pathway is executed for a patient (patient pathway) and information about each performed step is recorded (step of patient pathway). Finally, each performed step may be linked to a service that is executed for the patient (services performed) (Mans, van der Aalst and Vanwersch, 2015).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. The UML classes used to describe clinical pathways in the HRM. The original figure in (Mans, van der Aalst and Vanwersch, 2015, p. 49) also contained the ~150 attributes of those classes and provides cardinalities for the relations

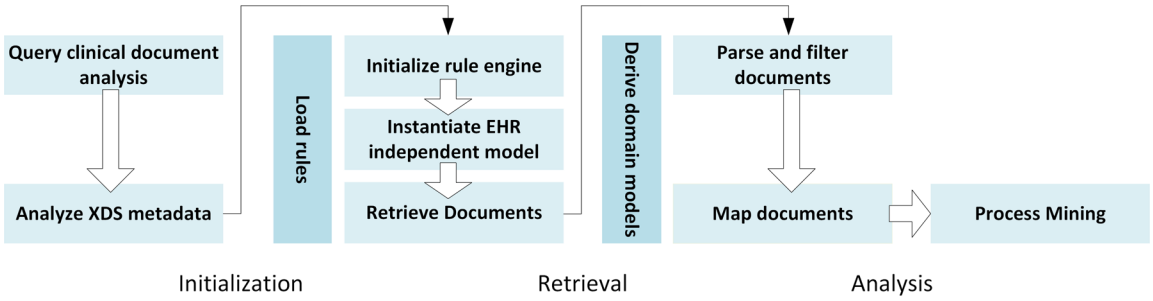


Using Clinical Document Architecture to Model Pathways

Integrating the Healthcare Enterprise uses existing standards such as CDA to enable healthcare interoperability. The IHE IT Infrastructure framework describes the profile Cross-Enterprise Document Sharing (XDS), which can be used to store and transmit medical documents. The XDS metadata allows to derive rules for creating an electronic health record (EHR) independent model (see Figure 2). The RIM model, which CDA is based on, can be used to develop a model for the healthcare data and documents. This model guarantees that data is structured in such a way that it is easily interpretable. This is necessary to develop a semantic model based on different data models. For example, for nursing care, it could describe the content of the care summary, its semantics and interpretation to guarantee semantic interoperability. This requires standardized structures for documents and a common language between the mentioned healthcare service levels, which is defined in documentation guidelines and schematron

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Modular, Rule-based Information Extraction Algorithm



rules. Provided terminologies can be accessed through a terminology server. In order to define key issues e.g. for the above-mentioned care summary example, the semantic model has to be developed in cooperation experts, in this case with healthcare personnel, favoring the use of internationally accepted and validated standards, for example. Thus, a basis for cross-enterprise exchange of care information is developed which provides the basics for the integration of healthcare service providers and process models.

Using Fast Healthcare Interoperability Resources to Model Pathways

The definition of clinical pathways in FHIR is modelled with the *PlanDefinition* resource, which defines a goal that shall be achieved and a set of actions to achieve that goal. Actions have a timing when they can/should occur, as well as trigger definitions and conditions if and when they are applicable. Each action has participants, explaining who is involved in the action and relations to other actions (HL7 2018), which allows the definition of highly complex clinical pathways. The FHIR *PlanDefinition* resource represents only an abstract concept of what should happen in a medical pathway and does not relate to a specific context such as actual patients or groups of persons.

To specify context-specific clinical pathways, such as one real patient to be treated, *CarePlan* resources can be used, which is based on a PlanDefinition, but can modify defined steps as needed for patient-specific treatment. Additionally, it documents activities, e.g. the steps performed during execution of the CarePlan and accompanying treatments. Each activity includes a template of what is going to happen (resource draft), or alternatively contains the resource that documents what happened. An example is a *MedicationRequest* documenting the administration of a medication, or an *Appointment* for the next treatment. Unlike the *PlanDefinition*, the steps in a *CarePlan* have no relation to each other (HL7, 2018).

Using FHIR, security auditing can be specified, which can be used for a fine-grained view of a medical pathway. The *AuditEvent* resource documents every single access to a FHIR resource according to the Five Ws (Who, What, When, Where, Why). Additionally, the *Provenance* resource tracks similar information when a resource is being modified (HL7, 2018).

Comparison

All three models support modelling clinical pathways. The HRM provides clearly defined classes to enable the modelling. CDA is based on the RIM and provides a document structure that enables mapping of documents to mine patient pathways from them. HRM and FHIR have clearly defined structures that make

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

modelling of pathways possible in one structured overview. In comparison, the view of CDA is distributed over several documents which need to be collected into one. CDA and FHIR enable semantic interoperability by using codings and value sets that define these codings. Such a mechanism is lacking in HRM. Both HRM and FHIR contain methods of logging an executed process, and allow the modelling of such logs. CDA itself is missing this functionality, but is provided similar mechanisms through IHE Auditing.

Interoperable ETL Approaches

To achieve high quality healthcare over institutional borders, all three levels of interoperability defined in the background, are needed between different healthcare providers. Based on existing use cases and using the methods for modelling clinical pathways, the HRM approach, the CDA approach and the FHIR approach are analyzed concerning semantic interoperability, especially how they add semantic information to the models, and process interoperability with focus on whether they support the ETL process. A detailed overview can be found in Table 2.

Analysis of the HRM

The analysis of the HRM is based on six use cases listed by (Mans, van der Aalst and Vanwersch 2015), where data described via the HRM was gathered from hospitals and used for process mining tasks like model discovery, conformance checking, bottleneck analysis and comparative analysis.

ETL Using the HRM

According to (Mans, van der Aalst and Vanwersch 2015), the HRM is instrumental in locating the data needed and facilitating the actual extraction. It supports the ETL process by giving the analysts an idea about what kind of process-related data can be found in a HIS.

The abstract HRM does not directly refer to the database structure and thus does not support the operation of *Extracting* of data. The *Transform* and *Load* steps of the ETL process are not described (Mans, van der Aalst and Vanwersch 2015).

Semantics in the HRM

The HRM model does not explicitly refer to code lists, structured value sets or nomenclatures. Attributes listed in the UML classes are identified by name, the semantics for those attributes are not provided in the textual description of the models (Mans, van der Aalst and Vanwersch 2015).

Analysis of HL7 FHIR

The FHIR approach was analyzed using a system for Multidisciplinary Team Meetings (MDTM) as described in (Krauss et al., 2017). MDTM are modeled in HL7 FHIR using the *PlanDefinition* resource. The models are then automatically transformed into the Business Process Model Notation (BPMN) and executed with a Workflow Engine. The process is documented using the *CarePlan* resources as well as FHIR *Auditing*. An ETL process, as described in the preceding section, could generate XES Event Logs, subsequently enabling Process Mining.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ETL Using FHIR

FHIR implicitly enables the entire ETL process. The *Extraction* from a data source is a given for any healthcare system implementing the FHIR standard, as the FHIR API allows reads and searches on resources.

The *Transformation* process can be executed directly in FHIR. In addition to the *ConceptMap* which allows semantic transformations, the *StructureMap* allows the mapping of any FHIR Resource to (or from) any other *Concept* defined by a *StructureDefinition*. This allows the implementer the definition of *StructureDefinitions* tailored to the exact requirements of the process-mining target. From there the *StructureMap* can describe how a regular FHIR resource can be transformed into the required structure – e.g. the eXtensible Event Stream (XES) used in recent process mining tools (Verbeek et al. 2010). Since the *StructureMap* contains machine-executable rules for transformation this mapping can be done fully-automated in a standardized way.

The *Loading* of the transformed resources can be achieved in several different ways. For one the FHIR API once again allows the access to the transformed resources similar to the *Extraction* process. Alternatively, the data-mining application can use the FHIR *Subscription* service to receive push messages while the process is executed. In another push-based approach one can create a FHIR *Operation* that publishes the transformed resources directly to the data mining application.

Semantics in the FHIR Model

Machine readable semantics of FHIR resources are handled through *Codings* that represent fields of resources, similar to *Codings* in the HL7 RIM. Each *Coding* consists of a *System* that defines the Code, the actual *Code* for machine-readability, and a *DisplayName* for human-readability. HL7 FHIR uses LOINC, SNOMED CT, HL7v3, ICD-10, and DICOM among others. (HL7 2018)

FHIR enables the restriction of allowed values in a *Coding*, and often provides default sets, or enforces usage of a specific set. This is documented in the *ValueSet* resource which groups *Codings* into a set. To enable semantic interoperability between different *Codings* the *ConceptMap* Resource exists to define unidirectional mappings between *ValueSets*. In addition, FHIR defines a *Terminology Service* specification that uses these resources to enable usage and transformation between *ValueSets* (HL7 2018).

Comparison of the Interoperability Status

The described models allow the interaction of technical components and systems. Further, they enable a larger interconnected system capability that transcends the local perspective of each participating subsystem, which complies with the technical interoperability defined by (Serrano et al., 2015). Besides technical standards, agreements are essential, how medical and domain specific terminologies are used, which have to be maintained and further developed over time. This way it is possible to relate pathways across various healthcare institutions, while preserving the intended meaning, which conforms to the definition of semantic interoperability in (Heitmann and Gobrecht, 2009). Thus, the analysis of non- and semi-structured health information should be based on a semi-automated approach, which allows semantic enrichment of the extracted and transformed models and provides semantic services to access terminology descriptions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Evaluation of model approaches based on interoperability levels

	HRM	FHIR	CDA
Use Case	Evaluation is based on use cases listed by (Mans, van der Aalst and Vanwersch 2015).	Multidisciplinary Team Meetings (MDTM) as described in (Krauss et al. 2017).	Based on the various definitions by IHE Technical Infrastructure
ETL	Supports parts of the ETL process. It does not support the operation of Extracting of data. The Transform and Load steps of the ETL process are not described (Mans, van der Aalst and Vanwersch 2015).	FHIR implicitly enables the entire ETL process.	Not supported in CDA as it only defines the data structure. Entire ETL enabled by IHE Profiles.
Semantics	No specification of semantics	Semantics specified through codings. ValueSet defines valid codings. ConceptMap enables unidirectional mapping between Valuesets Terminology Service enables usage and transformation between ValueSets.	Semantics specified through codings. ValueSet defines valid codings.
Interoperability	Complies to the technical interoperability defined by (Serrano et al. 2015).	Conforms to the definition of semantic interoperability in (Heitmann and Gobrecht 2009).	Enables technical and semantic interoperability due to implementation guides.
Automatic semantic interoperability	No since no form of coding exists.	Yes, through Terminology services. Process-Interoperability also possible with PlanDefinition	Yes.

Since the HRM model does not explicitly refer to code lists and semantic details are neither provided in the model definition nor description, semantic interoperability can hardly be achieved in an automated way using the HRM.

For IHE and CDA-based health data communication, XDS metadata as well as implementation guides and derived schematron rules are used to extract the clinical document types and which terminologies, coding sets and concepts are provided in the exchanged clinical information. Without a user-centred approach and the consideration of processes, semantic interoperability is not possible. Vice versa, knowledge about processes can only be extracted using semantic information provided by models. Thus, the information models are semantically enriched through aggregation and conversion of select information, profiles and codes based on the structure and requirements of the target domain. This not only allows the mapping of models from one domain to another (semantic) domain, which is partially based on (XSLT) templates, but further enriches derived process models and descriptions.

Thus, it also enables a sound interpretation of the content and enables to compare different data, find contradictory or incomplete data, and complete or correct the data. The models can also be used to check the plausibility of various values if the model supports similarity values. Through IHE compliance, the models provide interoperability across healthcare enterprises. Healthcare data can be exchanged electronically through the independent model, and the data can efficiently be combined/integrated and interpreted in context. Furthermore, these models allow the easy integration of further systems, services and processes.

For FHIR-based health data exchange, the operations framework is used to get definitions for a concept / find a concept within a ValueSet. Semantic interoperability is achieved using common information models and the terminology service, thus enabling process definitions in a certain domain as well

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

as across various domains or communities. The FHIR resources described above allow the definition (*PlanDefinition*, *CarePlan*) as well as the documentation (coarse granularity: *CarePlan*; fine granularity: *AuditEvent*, *Provenance*) of a medical pathway. This also allows a comparison between what should happen versus what actually did happen, which can be used to further process interoperability as well as check, compare and evaluate clinical pathways in and across institutions.

CONCLUSION

This article compares healthcare data models, which facilitate process mining, on a conceptual basis and analyzes their ability to describe clinical and patient pathways. Models resulting of standardized approaches are, although partially complex, easier to analyze than proprietary models, as information on standardized approaches is publicly available and overcome the limitations provided by approaches based on singular (hospital) information systems (Tang et al. 2006). Furthermore, it describes how the model and their design approaches support the ETL process to prepare data from different sources for subsequent process mining. By analyzing the different approaches and their implications regarding the support of the ETL process, the use cases for the models become apparent. In the Epilogue of (Mans, van der Aalst and Vanwersch 2015) the authors highlight the importance of the HRM to reason about questions that may or may not be answered using process mining. Moreover the HRM produces awareness of the data present in a hospital.

It is concluded that a standards-based top-down approach, and the derived FHIR approach respectively, is able to provide similar insights and, on top of that, operational support for the ETL process on all interoperability levels. As described above, the main reason for that is the ability to (automatically) integrate data from different sources by taking their syntactic and semantic properties into account. The proposed standardized approaches enable the analysis of even complex processes, allowing mining results like the causes of the increased length of stay, the lack of some essential care interventions in the treatment, the problems in following clinical guidelines, the discovery new care pathways, the discovery of best practices, the anomalies and the exceptions which may exist in processes, providing a better understood where to take action to improve the healthcare processes.

In conclusion all three analyzed models - HRM, CDA and FHIR – allow the modelling of clinical pathways. HRM and FHIR have specific features for this, and CDA contains metadata that allows the reconstruction of such pathways. HRM partially supports the ETL process, with the exception of data exctraction. FHIR, and CDA through the use of IHE Profiles, support the full ETL process. FHIR and CDA contain enough information to enable automatic semantic interoperability. This is not provided by HRM.

All three models have their merit for process mining, and clear strengths and weaknesses depending on the specific Use-Case process mining is applied in:

- HRM is weak when requiring semantic interoperability. But its connection to various proprietary solutions, and clear modelling of pathways make it useful when dealing with non-standardized systems.
- CDA has a clear advantage over HRM concerning semantic interoperability. Its major disadvantage is that it does not contain a definition for clinical pathways itself, and so while the pathways can be extracted from CDA a different representation, or alternatively a HL7 RIM template, is required.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- FHIR combines both advantages of HRM and CDA by being semantically interoperable and defining clear structures both for defining clinical pathways, as well as tracking them through execution. Its disadvantage over CDA is that, at the time of writing, IHE has little FHIR integration, making FHIR less widely used. The disadvantage over HRM is dealing with different standards or proprietary models. FHIR contains a Mapping Definition and Language to other models (not necessarily standardized!) but at the time of writing no publicly available implementation that enables automated transformations based on this definition exists.

Further research on the implementation of archetype-based systems should be conducted to complete the variety of data sources.

ACKNOWLEDGMENT

Parts of this research were supported by the Austrian Research Promotion Agency (FFG). Any opinions, findings and conclusions or recommendations in this paper are those of the authors and do not necessarily represent the views of the research sponsors.

REFERENCES

Benson, T., & Grieve, G. (2016). *Principles of Health Interoperability: SNOMED CT, HL7 and FHIR*. Springer International Publishing. doi:10.1007/978-3-319-30370-3

Chelsom, J., ... 2015. Document-Driven Care Pathways Using HL7 CDA. In *eTELEMED 2015: The Seventh International Conference on eHealth, Telemedicine, and Social Medicine*, Lisbon, Portugal.

Franz, B., & Mayr, H. (2011).Providing technical and semantic interoperability for integrated healthcare using an IHE-compliant system. In *Proceedings of the IADIS International Conference e-Health 2011*, 20-22 July, Rome, Italy.

HL7. (2016). HL7 Version 3 Domain Analysis Model: Care Plan. *Release*, *1*, HL7.

HL7. 2018. FHIR Release 3 (STU) Resource Index. Retrieved from http://hl7.org/fhir/resourcelist.html

Heitmann, K. U., & Gobrecht, K. (2009). *HL7 Communincation standards for the healthcare system*. Köln: LUP AG Lithographie & Printproduktion. (in German)

Integrating the Healthcare Enterprise. (n.d.). Retrieved from http://www.ihe.net

Krauss, O., Holzer, K., Schuler, A., Egelkraut, R., & Franz, B. (2017). Challenges and approaches to make multidisciplinary team meetings interoperable – the KIMBo project. *Studies in Health Technology and Informatics*, *236*, 63–69. PMID:28508780

Kühne, T. (2006). Matters of (Meta-)Modeling. Journal on Software and Systems Modeling, 5(4), 369-385.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mans, R., Schonenberg, H., Leonardi, G., Panzarasa, S., Cavallini, A., Quaglini, S., & van der Aalst, W. (2008). Process mining techniques: An application to stroke care. *Studies in Health Technology and Informatics*, *136*, 573. PMID:18487792

Mans, R., van der Aalst, W., & Vanwersch, R. J. (2015). *Process Mining in Healthcare: evaluating and exploiting operational healthcare processes*. Springer. doi:10.1007/978-3-319-16071-9

Norgall, T. (2003). Communication standard for health telematics: status and outlook. In Proceedings e-Health 2003. (in German)

Panella, M., Marchisio, S., & Di Stanislao, F. (2003). Reducing clinical variations with clinical pathways: Do pathways work? *International Journal for Quality in Health Care*, *15*(6), 509–521. doi:10.1093/intqhc/mzg057 PMID:14660534

Partington, A., Wynn, M., Suriadi, S., Ouyang, C., & Karnon, J. (2015). Process mining for clinical processes: A comparative analysis of four Australian hospitals. *ACM Transactions on Management Information Systems*, *5*(4), 19. doi:10.1145/2629446

Serrano, M. (2015). *Semantic Interoperability Release 2.0, AIOTI WG03 - IoT Standardisation*. Alliance for Internet of Things Innovation.

Sunyaev, A., Schweiger, A., Leimeister, L. M., & Krcmar, H. (2008). Software Agents for integration of information systems in healthcare. In Proceedings of e-Health 2008 (pp. 1455-1466). (in German)

Tang, P. C., Ash, J. S., Bates, D. W., Overhage, J. M., & Zands, D. Z. (2006, March 1). Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*, *13*(2), 121–126. doi:10.1197/jamia.M2025 PMID:16357345

Van der Aalst, W. M. (2011). Getting the Data. In *Process Mining* (pp. 95–123). Springer. doi:10.1007/978-3-642-19345-3_4

Verbeek, H. M. W., Buijs, J. C., Van Dongen, B. F., & van Der Aalst, W. M. (2010). XES, XESame, and ProM 6. *CAiSE Forum*, *72*, 60-75.

ENDNOTE

[1] i.s.h.med is a hospital information system developed by the Cerner Corporation

This research was previously published in the International Journal of Privacy and Health Information Management (IJPHIM), 6(2); edited by Francesco Longo and Letizia Nicoletti; pages 1-12; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 24
A Secure Robust and Privacy Enhanced Mobile Healthcare Framework

Hisham M. Alsaghier
Majmaah University, Kingdom of Saudi Arabia

Shaik Shakeel Ahamad
Majmaah University, Kingdom of Saudi Arabia

ABSTRACT

This article describes how the exponential growth of mobile applications has changed the way healthcare services function, and mobile healthcare using the Cloud is the most promising technology for healthcare industry. The mobile healthcare industry is in a continuous transition phase that requires continual innovation. There has been identified some of the challenges in the area of security protocols for mobile health systems which still need to be addressed in the future to enable cost-effective, secure and robust mobile health systems. This article addresses these challenges by proposing a secure robust and privacy-enhanced mobile healthcare framework (SRPF) by adopting a Community Cloud (CC), WPKI cryptosystems, Universal Integrated Circuit Cards (UICCs) and a Trusted Platform Module (TPM). All the security properties are provided within this framework. SRPF overcomes replay attacks, Man in the Middle (MITM) Attacks, Impersonation attacks and Multi-Protocol attacks as SRPF was successfully verified using a scyther tool and by BAN logic.

1. INTRODUCTION

Mobile cloud computing (MCC) is a new technology based on mobile web services. MCC is likely to play a vital role in mobile healthcare. Mobile cloud computing (MCC) is an integration of cloud computing (CC) into the mobile healthcare applications of mobile devices. In the cloud healthcare data can be stored and transmitted to hospitals and doctors, but security and privacy of this information is the major drawback and bottleneck of healthcare systems so in order to overcome these flaws US govern-

DOI: 10.4018/978-1-7998-1204-3.ch024

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ment introduced Health Insurance Portability and Accountability Act (HIPAA) to regulate US healthcare industry. Following are the motivations for adopting MCC in healthcare. MCC can improve healthcare services by providing Centralization, Performance, Modernization, Scalability, Portability, Collaboration and Virtualization. The main advantage of Mobile healthcare is it provides users easy and quick access to the healthcare infrastructures which includes PHR (patient health records) and doctors.

1.1. Motivations for the Research

1. According to ABI Research **a**ll the healthcare infrastructures are susceptible to attacks, in spite of this healthcare industry is spending very less amount for cybersecurity **(**Steven C. Morgan, 2015);
2. By 2020, 80 percent of healthcare data will be at cloud at some point of time (Julie Bird, 2016).

1.2. Related Work

(Mohammad Wazid, Sherali Zeadally, Ashok Kumar Das & Vanga Odelu, 2016)

Paper identifies some challenges in the realm of security protocols for mobile health systems which needs to be addressed in the future in order to ensure cost-effective, secure and robust mobile health systems. Following are the challenges outlined in this research work.

Challenge 1: Confidentiality, data integrity, accountability, availability, and access control are the most important security requirements for mobile health systems. For ensuring these security requirements, developing efficient key distribution protocols becomes challenging task in the mobile health system.

Challenge 2: Recent studies in the literature have shown that the public key operations (for example, elliptic curve cryptography) are practical in mobile devices. However, the Digital signature generation is very costly because of their computational complexity. Thus, efficiency of these credentials needs to be addressed. Since mobile healthcare applications deal with sensitive patient data, authenticity of the public keys should be efficient and cost effective to protect the data from unauthorized access.

Challenge 3: In contrast to public-key cryptography, symmetric key cryptography is superior and is easier to implement in term of its computational efficiency. However, symmetric key cryptography is not suitable because it relies on distributing the key in the mobile health system to provide different security services which include privacy of credentials, mutual authentication, and secrecy of session keys. Hence, designing efficient and flexible key distribution protocols for mobile healthcare applications needs to be addressed in the future.

Challenge 4: Obtaining passwords from unconscious patients is not possible so biometric mechanisms should be used for authentication. However, the biometric methods that work with unique biometric features from unconscious patients for identification purposes still needs further research attention in order to correctly authenticate an unconscious patient.

Challenge 5: The privacy of the information stored in the mobile healthcare application must be guaranteed. In this case, designing lightweight, efficient, and robust privacy enhancing techniques for the wearable sensors remains an area of future research.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1.3. Contributions Made

Secure Robust and Privacy enhanced Mobile Healthcare Framework based on Secure Element and Community Cloud: Formally Verified by BAN Logic and Scyther Tool.

Following are the contributions made by us:

1. We propose a Secure Robust and Privacy enhanced Mobile Healthcare Framework based on MHCC (Mobile Healthcare Community Cloud), CACC (Certifying Authority Community Cloud), UICC (Secure Element for both patient & Doctor) WPKI cryptosystems and Trusted Platform Module (TPM) at the Community Clouds;
2. We propose a procedure for personalizing mobile healthcare application at the patient and doctor side and a protocol in the proposed SRPF framework;
3. Our proposed framework consumes fewer resources as all the stakeholders are in the same Community Cloud (CC);
4. SRPF ensures Mutual authentication, data integrity, message confidentiality, Non-repudiation and Privacy (Patient) properties are ensured in this framework;
5. SRPF overcomes all the known attacks such as replay attacks, Man in the Middle (MITM) Attacks, Impersonation attacks and Multi-Protocol attacks as SRPF was successfully verified using scyther tool and BAN logic.

The rest of the paper is organized as follows: Section 2 proposes SRPF framework, Section 3 provides security analysis of the framework, Section 4 provides formal verification of SRPF protocol, Section 5 brings comparative analysis of the proposed framework with related work, Section 5 provides discussion, Section 6 concludes the work and the Appendix provides the SPDL code for SRPF and Results of SRPF in Simulation Environment provided by Scyther tool.

2. PROPOSED FRAMEWORK

2.1. Entities and Architecture of SRPF Framework

Our proposed secure mobile healthcare framework based on SRPF consists of the following participants:

1. **Patient (P):** Patient (P) is an important entity in the proposed framework possessing mobile phone with UICC;
2. **Trusted Service Manager (TSM):** Certifying Authority (CA) is responsible for issuing and authenticating of public keys. It issues certificates, binds public keys, manages serial numbers, and revokes certificates. Our Proposed framework allocates an additional responsibility of Trusted Service Manager (TSM) for CA. TSM is trusted by all the stakeholders in the ecosystem;
3. **Mobile Network Operator (MNO):** MNO is involved indirectly involved in this framework. It provides 4G connectivity and helps in enabling OTA (Over The Air) transaction. UICC of both the patient and doctor are issued by MNSP;
4. **Doctor:** Doctor is an important entity in our proposed SRPF framework;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. **Mobile Healthcare Community Cloud (MHCC):** MHCC caters the needs of Healthcare services to the citizens. It provides services to the patients by bringing all hospitals in the same cloud. It uses mobile web services to connect the mobile web using internet technology. MHCC will be managed by the central government of the country. Mobile Healthcare Community Cloud (MHCC) consists of Hardware Security Module HSM's of hospitals personalized by the respective hospitals. The connectivity between MHCC and Patients are through 4G/5G networks protected by TLS protocol, ECDSA and AES algorithms. Our proposed MHCC caters the needs of Healthcare services to the citizens. It provides services to the patients by bringing all hospitals in the same cloud.

2.1.1. Components at the MHCC

1. **UDDI Registry:** UDDI registry contains information about MHCC services and communication standards and interfaces used;
2. **Secure Token Service (STS) Server:** This entity authenticates patients and doctors credentials which are issued by TSM and issues a token which is used to carry out transactions. Communication Manager (ComM) acts as a catalyst between patients, doctors and MHCC by establishing a secure and reliable communication channel thereby ensuring reliable communication security;
3. **Time Stamping Authority (TSA):** It offers time stamping services for the messages generated in the MHCC;
4. **Synchronization Manager (SyM):** The main responsibilities of Synchronization Manager include personalization of mobile healthcare application of both patient and doctor. It also exchanges encrypted messages with both patient and doctor and finally it ensures scalability in the framework;
5. **Security and Auditing Manager (SAM):** The Security Manager is responsible for managing the credentials of all the stakeholders such as user name, passwords, biometric data, public keys and symmetric keys. The Security Manager verifies the certificates of all patients and doctors before exchanging messages with them. The Security Manager archives all the evidences in its database, evidences include digital signatures, cryptographic hash of password and encrypted messages using symmetric keys and finally Logs are backed up once in a month. In case of disputes among the stakeholders SAM can provide proof in the court of law with Long term validity preservation of Digital Signature and encrypted messages (using symmetric keys) even after completely destroying/losing all the keys and ECDSA algorithm once signed can prove that data is valid – by matching confirmation data with log data, time stamps, nonce and old certificates;
6. **Certifying Authority Community Cloud (CACC):** CACC caters the needs of both PKI and Wireless PKI services to its citizens. CACC will be managed by the central government of the country. It provides security services to the citizens by bringing all the stakeholders (RA, TSA & Directory of CA) in the same community cloud. CACC will be managed by CA under the directions of central government of the country. CACC consists of Hardware Security Module HSM's for all the stakeholders and these are personalized by CA only. The connectivity between CACC and Patients are through 4G/5G networks protected by TLS protocol, ECDSA and AES algorithms. CACC contains a Registration Authority (RA) for checking the credentials of the patients, Time Stamping Authority (TSA) provides time stamping services, OCSP which helps in checking the validity of the certificates and Directory of CA maintaining all the X.509 and short lived certificates.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.2. Procedure for the Personalization of MHA and TPM in Community Cloud

Step 1: If Patient (P) and Doctor (D) does not possess public key then Hospital (H) acts as a Registration Authority (RA) for both Patient (P) and Doctor (D) for issuing certificates. Certification Authority (CA) issues both X.509 and Short Lived Certificates (SLC) to all the stakeholders. RA checks the certificate of the SEs of both Patient (P) and Doctor (D) and maps the serial number and SE certificate to the Patient (P) and Doctor (D) national identity. All the stakeholders in the proposed framework generate their credentials in the tamper resistant hardware such as Secure Element (Patient (P) and Doctor (D)) and the Trusted Platform Module (TPM)/HSM of the Hospital in the community cloud.

Step 2: Trusted Platform Module (TPM) of the Hospital in the community cloud builds the database of the registered Patient (P) and Doctor (D).

Step 3: All the Patient (P) and Doctor (D) will be issued certificates and Patient (P) will be issued with a Traceable Anonymous Certificate (TAC) as per the RFC 5636 (S.Park, H.Park,Y.Won,J.Lee,Kisa & Kent, 2009) [14].

Step 4: Patient (P) and Doctor (D) will be asked to download mobile healthcare application which will be uploaded by the hospital in the community cloud, before downloading the mobile healthcare application Patient (P) and Doctor (D) will check the authenticity of the mobile healthcare application by downloading the certificate. If the checking is successful they accept the mobile healthcare application or report about its failure to the hospital.

Step 5: Patient (P) and Doctor (D) will check the TPM certificate belonging to the hospital which is in the community cloud. If the checks are successful they can start using Mobile healthcare application:

1. Patient (P) and Doctor (D) validates the platform certificate of TPM of the hospital which is in the community cloud by using the validation procedure given in D.R. Stinson 2006 for certificates;
2. Patient (P) and Doctor (D) delegates OCSP to validate Mobile Healthcare Application Certificate;
3. If the responses from OCSP is positive, Patient (P) and Doctor (D) requests Hospital TPM to download and install Mobile Healthcare Application on the SE. /* provisioning step */;
4. University TPM personalizes Mobile Healthcare Application which is in the SE's of Patient (P) and Doctor (D).

Step 6: SRPF personalizes mobile healthcare application of both the Patient and Doctor separately which is in their UICC of their respective applications using the protocol given in (Ahamad, Sastry & Udgata, 2014).

Step 7: SRPF allocates shared symmetric key for both the Patient and Doctor separately in their respective applications of UICC.

Step 8: MHCC sends the information of allocation of Doctor to the Patient (P) and sends Doctor's public key and Doctor's identity using the shared symmetric key between Patient and Hospital.

Step 9: Mobile healthcare application of the patient is personalized in such a manner that when the application senses abnormality in patient it sends message to both hospital and Doctor with the location of the patient.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.3. Proposed Protocol

We propose a secure mobile healthcare protocol (SRPF):

$$Step\ 1: MHCC \rightarrow D: \{MS1\}_{SYYKEY_{DMhcc}}$$
$$MS1: \{UN, PW, PID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}\}$$

Step 1: MHCC allocates Doctor to the Patient (P) and sends a message containing symmetric key shared between Patient and MHCC in order to decrypt the patient's encrypted files from the MHCC database in emergency. The message also contains Timestamp, Nonce, Username (UN), PW (Password) and PID (Patient Identity). UN and PW are needed to authenticate doctor into the MHCC database containing all the patients' records.

$$Step\ 2: D \rightarrow MHCC: \{MS2\}_{SYYKEY_{DMhcc}}$$
$$MS2: \{ACK, PID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}, T_D, N_D\}$$

Step 2: Doctor acknowledges (ACK) for receiving shared symmetric key between Patient and MHCC. Doctor encrypts the sent message with the shared symmetric key between Doctor and MHCC.

$$Step\ 3: MHCC \rightarrow P: \{MS3\}_{SYYKEY_{PMhcc}}$$
$$MS3: \{PID, DID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}\}$$

Step 3: After receiving acknowledgement (ACK) from Doctor, MHCC sends PID (Patient Identity), DID (Doctor's Identity) in step 3.

$$Step\ 4: P \rightarrow MHCC: \{MS4\}_{SYYKEY_{PMhcc}}$$
$$MS4: \{ACK, PID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}, T_P, N_P\}$$

Step 4: Patient acknowledges (ACK) for doctors information allocated to him in step 4.

3. SECURITY ANALYSIS

Our proposed protocol ensures all the security properties such as authentication, data integrity and non-repudiation as the messages are exchanged using Digital Signatures based on ECDSA and encrypted using AES algorithm. Our proposed protocol ensures as intruder fails in forging patient's signature because he will be not in possession of patient's private key. Confidentiality is ensured using shared symmetric keys. In our proposed SRPF framework privacy of the patient is ensured using Traceable Anonymous Certificate (TAC) as per the RFC 5636 (Park et al., 2009). This section provides theoretical analysis of our proposed protocol in order to prove that the proposed protocol withstands attacks on authenticity, confidentiality, integrity and non-repudiation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.1. Assumptions

We have made a number of practical and realistic assumptions in our proposed protocol. In the following, we first describe those assumptions.

Assumption 1: In our proposed framework Certifying Authority (CA) is a Trusted Service Manager (TSM) for all the stakeholders in the ecosystem and is fully trusted by all the stakeholders. Important duties of CA includes issuing certificates to all the stakeholders and Application testing and certification.

Assumption 2: Patient credentials were generated (by the patient) and they reside in the trusted part of tamper resistant Secure Element i.e. UICC of the patient and no one is authorized to access and manipulate these credentials as they are protected by password and biometric.

Assumption 3: In our proposed framework, symmetrically encrypted messages cannot be decrypted by any other stakeholders, unless they have the session/shared key used to encrypt/decrypt the message by the sender.

Assumption 4: In our proposed framework, messages are digitally signed by the sender using private key resided in the trusted part of the tamper resistant Secure Element (i.e. UICC) using ECDSA algorithm.

Assumption 5: SRPF framework adopts (TAC) given in RFC 5636 (Park et al., 2009) which offers privacy or anonymity to the Patient's identity, TAC is a X.509 certificate containing a pseudonym of Patient (P). CA has the ability to map TAC to the real identity of the Patient (P).

Assumption 6: Mobile Healthcare Application (MHA) of the patient and doctor is personalized by the Mobile Healthcare Community Cloud (MHCC) i.e. a symmetric key is agreed between MHCC and patient and MHCC and doctor.

3.2. Threat Model

We define our threat model as follows:

Definition 1 (Attack on authentication): A malicious intruder (In) is an intruder with the ability to monitor, intercept and introduce new stakeholders on behalf of genuine stakeholders.

Definition 2 (Attack on confidentiality or Information Privacy): An intruder (In) is an unauthorized entity who is capable of viewing the data before it reaches the receiver.

Definition 3 (Attack on integrity): An intruder (In) is an unauthorized entity who is capable of modifying the data before it reaches the receiver.

Definition 4 (Attack on Non-repudiation): See below:

1. Digital signatures generated in SRPF framework are called as Qualified Electronic Signatures because they are generated in Secure Signature Creation Device (SSCD);
2. A malicious intruder (In) attack on Non repudiation, in which an adversary is capable of generating Qualified Electronic Signature on behalf of any one or all the stakeholders (using private key) of SRPF framework.

Definition 5 (Attack on Anonymity of the patient): Doctor with malicious intension tries to know the real identity of Patient (P).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Definition 6 (Attack on Communication Privacy): A malicious intruder (In) tries to capture the data communicated during the transit.

3.3. Security Proof

In this sub-section, we prove our proposed framework (SRPF) is secure against the threat model given in B sub section of this section considering the assumptions explained in Section 1. This is achieved through seven theorems with corresponding proofs as follows:

Theorem 1: Mobile Healthcare Application (MHA) of the Patient (P) is personalized by the MHCC (i.e. Shared symmetric key is agreed and installed in the MHA).

Proof: Mobile Healthcare Application (MHA) of the Patient (P) is personalized by the MHCCC Server by following the procedure proposed in (Shaik Shakeel Ahamad, V.N. Sastry, Siba K. Udgata, 2014).

Theorem 2: An Intruder (In) cannot read and modify the secret information transmitted by the stakeholders during the transit.

Proof: In SRPF framework messages are transmitted by encrypting (AES algorithm) and digitally signing (ECDSA algorithm) by all the stakeholders.

Theorem 3: Doctor fails in getting the real identity of the patient.

Proof: Patient (P) gets anonymous identity from CA; Hospital and CA knows the real identity of P. CA issues TAC to the Patient (P), so Doctor will not be revealed with the original identity of Patient (P).

Theorem 4: SRPF consumes fewer resources from the patient's perspective.

Proof: SRPF framework uses digital signature mechanism based on ECDSA which consumes very less resources. Patient (P)'s interaction with the other stakeholders in the framework is very less thereby consuming fewer resources for the patient.

Theorem 5: SRPF framework ensures both communication security.

Proof: SRPF framework ensures communication security using SSL/TLS protocol.

Theorem 6: SRPF framework generates Qualified Electronic Signatures.

Proof: Secure Signature-Creation Device (SSCD) configures software or hardware or both in order to implement Signature-Creation Data (SCD). Signature-creation data (SCD) contains private keys used to generate digital signature or electronic signature. Stakeholders in SRPF framework generate digital signatures in UICC (which is a secure element) and in the Grid of Secure Elements (at the TSM and at MHCC) so they are considered as Qualified Electronic Signatures.

Note: UICC (which is a secure element) and Grid of Secure Elements are Secure Signature Creation Device (SSCD).

Theorem 7: SRPF framework withstands all the known attacks.

Proof: Our proposed framework can withstand all the attacks; SRPF uses timestamps and nonce in order to ensure the freshness of the messages which avoids replay attacks, SRPF overcomes impersonating attacks as the intruder does not possess P's private key and SRPF framework withstands Man In The Middle (MITM) Attack as the attacker do not possess MHCC/Patient/Doctor's private key.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. FORMAL VERIFICATION OF SRPF PROTOCOL

4.1. Formal Verification Using BAN Logic

A security protocol is a communication protocol which exchanges encrypted messages by using cryptographic mechanisms (Muhammad et al., 2006). Popular and carefully designed protocols were found out to have security breaches (Muhammad et al., 2006). We have analyzed the protocol using BAN logic ((Abadi et al., 1993) and (Burrows et al., 1990)).

4.2. Assumptions for the Analysis and Verification of the Proposed Protocol

4.2.1. Assumptions About Keys and Secrets

'S' is a set of stakeholders containing {P, MHCC, D and CA}. These assumptions gives a brief overview of public and private keys possessed by all the stakeholders. CA certifies all the certificates and knows all the public keys of the stakeholders (AS1, AS2).

AS1. CA **believes** $\left(\forall S \in \{\text{P, MHCC, D and CA}\} \overset{K_s}{\mapsto} S \right)$. Certification Authority CA believes that all the stakeholders have their own public keys to communicate.

AS2. $S \in$ {P, MHCC, D and CA} believes $\overset{K_{ca}}{\mapsto} CA$. All the stakeholders in the framework knows the public key and certificate of the certification authority CA.

4.2.2. Assumptions About Freshness

Assumption **AS3** specifies freshness of quantities. For instance, if the Patient P sees quantity (N_p) n a message then the Patient P can conclude that it is a replay message.

AS3. P believes freshness (N_p), MHCC believes freshness (N_{Mhcc}), D believes freshness (N_d).

Every stakeholder believes nonce generated by him/her is fresh.
Assumption **AS4** is about validity time of certificates and timestamps which ensures timeliness.

AS4. TS_x TS_y are the timestamps generated by the stakeholders X and Y ({U, IB, AB, PG, CP, RI, and CA}) which ensures **timeliness** of the messages exchanged.

4.2.3. Assumptions About Trust

These assumptions give a brief overview of trust level on each stakeholder.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

AS5. ($\forall S, Q \in$ {P, MHCC, D and CA}. S **believes** CA **controls** $\overset{K_{ca}}{\mapsto} Q$. Every stakeholder trusts the Certification Authority CA.

AS6. $\forall$ belief X, CA **believes** (W **controls** (P **believes** X)). The Certification Authority CA trusts the Patient P that UICC or Secure Element (SE) (W) to relay Patient P's beliefs.

4.3. Verification of Our Proposed Protocol Using BAN Logic

$$Step\ 1: MHCC \rightarrow D: \{MS1\}_{SYYKEY_{DMhcc}}$$

$$MS1: \{UN, PW, PID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}\}$$

D decrypts the received $\{\boldsymbol{MS}1\}_{\boldsymbol{SYYKEY_{DMhcc}}}$ from the assumptions **AS1, AS2, AS5, AS6 & AS7**

D **believes** $\{\boldsymbol{MS}1\}_{\boldsymbol{SYYKEY_{DMhcc}}}$ statement (1)

D verifies the public key of MHCC (**AS7**) received from MHCC which mainly includes [9] (D.R. Stinson 2006).

If the verification of certificate is successful then:

D **believes MHCC said** $\{\boldsymbol{MS}1\}_{\boldsymbol{SYYKEY_{DMhcc}}}$ statement (2)

D **believes fresh** $\boldsymbol{T_{Mhcc}}$ from **AS3** statement (3)

D **believes fresh** $\boldsymbol{N_{Mhcc}}$ from **AS4** statement (4)

So from the statements 1 to 4:

D **believes** $\{\boldsymbol{MS}1\}_{\boldsymbol{SYYKEY_{DMhcc}}}$

$$Step\ 2: D \rightarrow MHCC: \{MS2\}_{SYYKEY_{DMhcc}}$$

$$MS2: \{ACK, PID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}, T_D, N_D\}$$

MHCC decrypts the received $\{\boldsymbol{MS}2\}_{\boldsymbol{SYYKEY_{DMhcc}}}$ **from D** from the assumptions of **AS1, AS2, AS5, AS6 & AS7**

MHCC **believes** $\{\boldsymbol{MS}2\}_{\boldsymbol{SYYKEY_{DMhcc}}}$ statement (5)

MHCC validates the public key of D (**AS7**) received from D which mainly includes [9] (D.R. Stinson 2006):

If the verification of certificate is successful then:

MHCC **believes D said** $\{\boldsymbol{MS}2\}_{\boldsymbol{SYYKEY_{DMhcc}}}$ statement (6)

MHCC believes **fresh** $\boldsymbol{T_D}$ from **AS3** statement (7)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MHCC **believes fresh** N_D from **AS4** statement (8)

So from the statements 5 to 8:

D **believes** $\{MS2\}_{SYYKEY_{DMhcc}}$

$$Step\ 3: MHCC \rightarrow P: \{MS3\}_{SYYKEY_{PMhcc}}$$
$$MS3: \{PID, DID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}\}$$

P decrypts the received $\{MS3\}_{SYYKEY_{PMhcc}}$ **from MHCC** from the assumptions of **AS1, AS2, AS5, AS6 & AS7**

D **believes** $\{MS3\}_{SYYKEY_{PMhcc}}$ statement (9)

P validates the public key of MHCC (**AS7**) received from MHCC which mainly includes (D.R. Stinson 2006).

If the verification of certificate is successful then:

P **believes MHCC said** $\{MS3\}_{SYYKEY_{PMhcc}}$ statement (10)

P **believes fresh** T_{Mhcc} from **AS3** statement (11)

P **believes fresh** N_{Mhcc} from **AS4** statement (12)

So from the statements 9 to 12:

P **believes** $\{MS3\}_{SYYKEY_{PMhcc}}$

$$Step\ 4: P \rightarrow MHCC: \{MS4\}_{SYYKEY_{PMhcc}}$$
$$MS4: \{ACK, PID, SYYKEY_{PMhcc}, T_{Mhcc}, N_{Mhcc}, T_P, N_P\}.$$

MHCC decrypts the received $\{MS4\}_{SYYKEY_{PMhcc}}$ from P from the assumptions of **AS1, AS2, AS5, AS6 & AS7**

MHCC **believes** $\{MS4\}_{SYYKEY_{PMhcc}}$ statement (13)

MHCC validates the public key of P (**AS7**) received from P which mainly includes (D.R. Stinson 2006).

If the verification of certificate is successful then:

MHCC **believes P said** $\{MS4\}_{SYYKEY_{PMhcc}}$ statement (14)

MHCC believes **fresh** T_P from **AS3** statement (15)

MHCC **believes fresh** N_P from **AS4** statement (16)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

So from the statements 13 to 16:

MHCC believes $\{MS4\}_{SYYKEY_{PMhcc}}$

4.4. Formal Verification Using Scyther Tool

Our proposed protocol is written in Security Protocol Description Language (.spdl) of the Scyther simulation environment. Scyther tool is used to verify the security protocols (security services provided in Table 1). We define the role of HC, D and P, where HC is the Healthcare Center, D is Doctor and P is the Patient for scythe tool. In our scenario, MC, D and P exchange information. We introduce five types of attacks. The first type of attack is defined for the transmission among MC, D and P (integrity), the second attack is defined where an adversary acquires the properties of any one of the three entities and sends the attack data packets to other entities involved in the framework (authentication) and the third attack is defined adversary try to read the data within interval *t* (Confidentiality). We have experimented with 100 numbers of runs for each claim to found out that there are no attacks in our proposed protocol, in addition to these, we follow the default properties (Table 2) of Scyther tool. Outcomes provided in Table 3.

Table 1. Mapping between security objectives and security properties

Claims of Security	Scyther Security Services that Verify Claims
Authenticity	Alive, WeakAgree, NiAgree
Detection of replay, reflection and MITM attack	NiSynch
Confidentiality of the message	claim_HC1 (HC, Secret, Khcd); claim_HC2 (HC, Secret, nhc); claim_HC3 (HC, Secret, np);
Non repudiation	Agreement (NiAgree, NiSynch) between the roles, Aliveness and Authenticity using WeakAgree

Table 2. Setting parameters for the Scyther verification tool

Verification Parameters		Advanced Parameters	
Title	**Value**	**Title**	**Value**
Maximum number of runs	100	Search Pruning	Find all attacks
Matching Type	Typed matching	Maximum number of patterns per claim	100

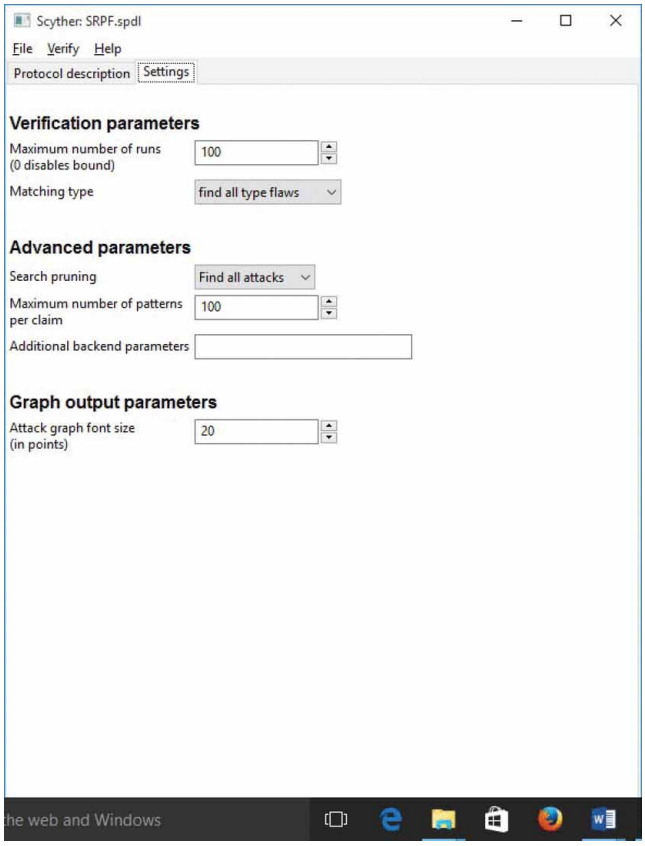
Table 3. Summarized outcome of automated security claims using Scyther verification tool

ROLES	Weakagree	Secrecy	Alive	Nisynch	Niagree	Comment
Patient (P)	Pass	Pass	Pass	Pass	Pass	Pass
Doctor (D)						
HC						

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Attack model:** In our framework we focus on integrity attacks, confidentiality attacks and authentication attacks as discussed above. In integrity attacks, an attacker observes the encrypted data travelling on the network that contain information. In confidentiality attack, attacker continuously observe the data flow and try to read the data. In authentication attacks, an attacker can observe a source node, and try to get the behavior of the source node. We assume that he/she is able to determine the source node's behavior. In such cases, the attacker can introduce an authenticated node and act as the original source node. In our concept, we are using trusted modules in sensors to store the secret information and procedure for key generation and encryption (such as TPM);
- **Results:** Our proposed protocol SRPF is written in SPDL (Security Protocol Description Language) for simulating in Scyther environment. Scyther (Cremers. 2006) & (Cremers & Lafourcade, 2007) is a tool which is used to verify, falsify and analyze security protocols. Scyther (Cremers. 2006) & (Cremers & Lafourcade, 2007) has many unique features which make it popular for analyzing infinite sets of traces in terms of patterns, unbounded verification with guaranteed termination and finally analyzing multi-protocols. Box 1 in the Appendix presents the code of SRPF in SPDL with results. The result window in Figure 1 of the Appendix shows simulation environment of scyther tool with its parameters. The result window in Figure 2 of the Appendix shows the "Verification Claim" Procedure of Scyther Tool and the result window in Figure 3 of the Appendix shows the "Automatic Claim" Procedure of Scyther Tool.

Figure 1. Scyther simulation environment with its parameters



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Result of SRPF protocol using "Verification Claim" procedure of Scyther tool

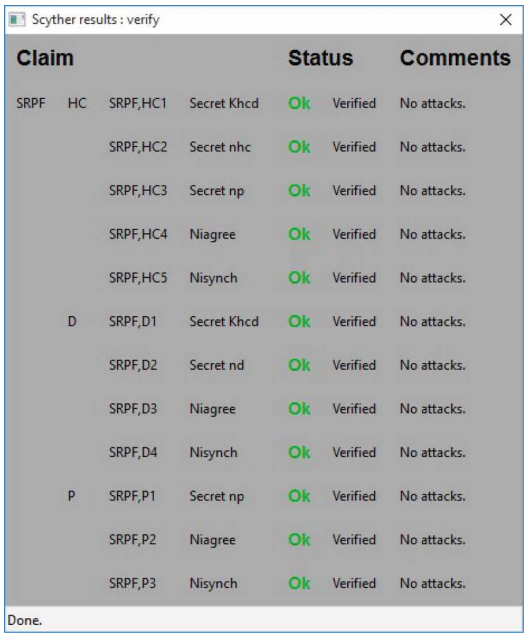
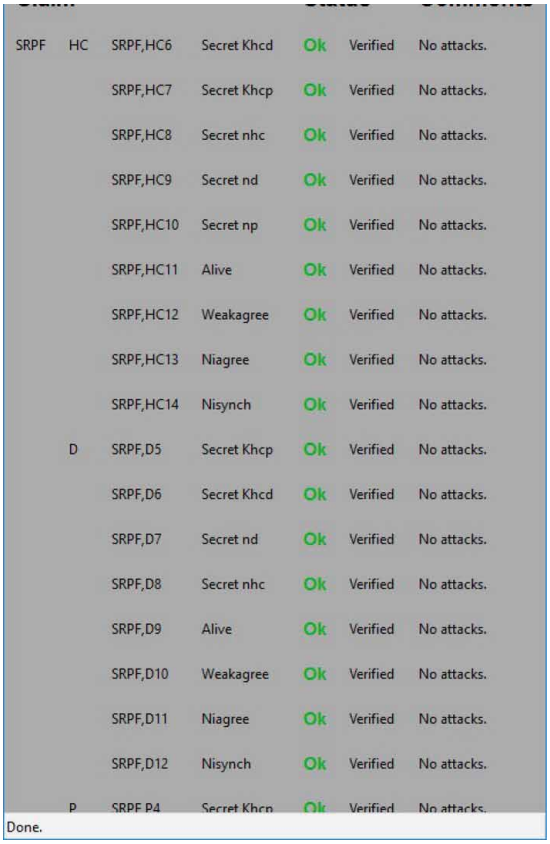


Figure 3. Result of SRPF protocol using "Automatic Claim" procedure of Scyther tool



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. COMPARATIVE ANALYSIS OF SRPF WITH RELATED WORK

Results in Table 4 are as follows. Yes: the protocol protects against that attack or provides that feature; No: the protocol does not protect against that attack or does not provide that feature. N.A.: not applicable

6. DISCUSSION

Following are the remedies provided by SRPF to challenges outlined in (Wazid et al, 2016) for enabling cost-effective, secure and robust mobile health systems.

6.1. Remedy to Challenge 1

Mobile Healthcare Application (MHA) of the patient and doctor is personalized by the Mobile Healthcare Community Cloud (MHCC) i.e. a symmetric key is agreed between MHCC and patient and MHCC and doctor. In SRPF symmetrically encrypted messages cannot be decrypted by any other stakeholders, unless they have the session/shared key used to encrypt the message by the sender, messages are digitally signed by the sender using private key resided in the trusted part of the tamper resistant Secure Element (i.e. UICC) using ECDSA algorithm. So all the security properties such as confidentiality, mutual authentication, data integrity, non-repudiation and access control are ensured using efficient key distribution protocol in the proposed framework.

Table 4. Comparative analysis of SRPF with related work

Protocols Functionality Features	(Arshad et al., 2015)	(Mishra et al., 2014)	(Mir, et al., 2015)	(Das et al., 2015)	(Siddiqui et al., 2013)	(Das, 2015)	(David, 2016)	(Moon et al., 2016)	(Mir et al., 2015)	SRPF
Ensures Confidentiality	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ensures Data Integrity	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Ensures Non-Repudiation	No	Yes	Yes	Yes	No	No	No	No	No	Yes
Ensures two factor authentication	Yes	No	No	No	No	No	No	No	Yes	Yes
Ensures three-factor authentication	No	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Ensures Patient's anonymity	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes
Ensures mutual authentication	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Ensures session key agreement	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Withstands Replay attacks	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Withstands Impersonation attacks	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Withstands Man In The Middle Attacks	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Withstands Multi-Protocol Attacks	No	No	No	No	No	No	No	No	No	Yes
Withstands stolen smart card/mobile device attack	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Formal Verification of Protocol	No	No	No	No	No	No	No	No	No	Yes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6.2. Remedy to Challenge 2

In our proposed framework SRPF messages are digitally signed by the sender using private key resided in the trusted part of the tamper resistant Secure Element (i.e. UICC) using ECDSA algorithm. Patient credentials were generated (by the patient) and they reside in the trusted part of tamper resistant Secure Element i.e. UICC of the patient and no one is authorized to access and manipulate these credentials as they are protected by password and biometric. Mobile Healthcare Application (MHA) of the patient and doctor is personalized by the Mobile Healthcare Community Cloud (MHCC) i.e. a symmetric key is agreed between MHCC and patient and MHCC and doctor. So the authenticity of the public keys are efficient and cost effective as they protect data from unauthorized access because of the personalization (for private key) of UICC (of the patient and doctor) and TPM (for private key at the MHCC) by the patient, doctor and by the Hospital.

6.3. Remedy to Challenge 3

Mobile Healthcare Application (MHA) of the patient and doctor is personalized by the Mobile Healthcare Community Cloud (MHCC) i.e. a symmetric key is agreed between MHCC and patient and MHCC and doctor thereby achieving credentials privacy, mutual authentication, and session key security using symmetric key cryptography.

6.4. Remedy to Challenge 4

Patient credentials were generated (by the patient) and they reside in the trusted part of tamper resistant Secure Element i.e. UICC of the patient and no one is authorized to access and manipulate these credentials as they are protected by fingerprint biometric. So our proposed SRPF framework uses fingerprint biometric that has unique biometric features as it works in unconscious patients for identification purposes.

6.5. Remedy to Challenge 5

In our proposed framework, Mobile Healthcare Application (MHA) of the patient and doctor is personalized by the Mobile Healthcare Community Cloud (MHCC) i.e. a symmetric key is agreed between MHCC and patient and MHCC and doctor. So symmetrically encrypted messages cannot be decrypted by any other stakeholders, unless they have the session/shared key which is used to encrypt the message by the sender.

7. CONCLUSION

In this article, we propose a Secure Robust and Privacy enhanced Mobile Healthcare Framework (SRPF) which addresses the challenges outlined in the literature by adopting Community Cloud (CC), WPKI cryptosystems, Universal Integrated Circuit Card (UICC) and Trusted Platform Module (TPM). All the security properties such as Mutual Authentication, Integrity of the message, Privacy of the messages exchanged, Non-repudiation and Privacy of the Patient are ensured in this framework. SRPF overcomes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

all the known attacks such as replay attacks, Man in the Middle (MITM) Attacks, Impersonation attacks and Multi-Protocol attacks as SRPF was successfully verified using scyther tool and by BAN logic.

ACKNOWLEDGMENT

The authors would like to thank Deanship of Scientific Research at Majmaah University for supporting this work under Project Number 37/111.

REFERENCES

Abadi, M., Burrows, M., Kaufman, C., & Lampson, B. (1993). Authentication and delegation with smart-cards. *Science of Computer Programming*, *21*(2), 93–113. doi:10.1016/0167-6423(93)90002-7

Ahamad, S. S., Sastry, V. N., & Udgata, S. K. (2014). Secure mobile payment framework based on UICC with formal verification. *IJCSE*, *9*(4), 355-370.

Ali, M.: Green. 2009. cloud on the horizon. In *Proceedings of the 1st International Conference on Cloud Computing (CloudCom)*, Manila (pp. 451–459). 10.1007/978-3-642-10665-1_41

Arshad, H., Teymoori, V., Nikooghadam, M., & Abbassi, H. (2015). On the security of a two-factor authentication and key agreement scheme for telecare medicine information systems. *Journal of Medical Systems*, *39*(8). doi:10.100710916-015-0259-6 PMID:26084586

Burrows, M., Abadi, M., & Needham, R. (1990). A logic of authentication. *ACM Transactions on Computer Systems*, *8*(1), 18–36. doi:10.1145/77648.77649

Cremers, C., & Lafourcade, P. Comparing State Spaces in Automatic Security Protocol Verification. *Proceedings of the 7th International Work-shop on Automated Verification of Critical Systems (AVoCS"07)* (pp. 49 – 63).

Cremers, C. J. F. (2006). Scyther-Semantics and Verification of Security Protocols [Ph.D. Dissertation]. Eindhoven University of Technology.

Cybersecurityventures. (n.d.). Cybersecurity market report. Retrieved from http://cybersecurityventures.com/cybersecurity-market-report-q2-2015/

Das, A. K. (2015). A secure user anonymity-preserving three-factor remote user authentication scheme for the telecare medicine information systems. *Journal of Medical Systems*, *39*(3), 1–20. doi:10.100710916-015-0218-2 PMID:25677956

Das, A. K., Odelu, V., & Goswami, A. (2015). A secure and robust user authenticated key agreement scheme for hierarchical multimedical server environment in TMIS. *Journal of Medical Systems*, *39*(9). doi:10.100710916-015-0276-5 PMID:26245854

David, D. B. (2016). Mutual authentication scheme for multimedia medical information systems. *Multimed. Tools Appl*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fiercehealthcare. (2014, November 21). 2015 healthcare predicitions: growth in analytics, mobile, security risks. Retrieved from http://www.fiercehealthcare.com/it/2015-healthcare-predictions-growth-analytics-mobile-security-risks

Mir, O., & Nikooghadam, M. (2015). A secure biometrics based authentication with key agreement scheme in telemedicine networks for e-health services. *Wireless Personal Communications*, *83*(4), 2439–2461. doi:10.100711277-015-2538-4

Mir, O., van der Weide, T., & Lee, C. C. (2015). A secure user anonymity and authentication scheme using AVISPA for telecare medical information systems. *Journal of Medical Systems*, *39*(9), 1–16. doi:10.100710916-015-0265-8 PMID:26242748

Mishra, D., Mukhopadhyay, S., Kumari, S., Khan, M. K., & Chaturvedi, A. (2014). Security enhancement of a biometric based authentication scheme for telecare medicine information systems with nonce. *Journal of Medical Systems*, *38*(5), 1–11. doi:10.100710916-014-0041-1 PMID:24771484

Moon, J., Choi, Y., Kim, J., & Won, D. (2016). An improvement of robust and efficient biometrics based password authentication scheme for telecare medicine information systems using extended chaotic maps. *Journal of Medical Systems*, *40*(3). doi:10.100710916-015-0422-0 PMID:26743628

Muhammad, S., Furqan, Z., & Guha, R. K. (2006). Understanding the intruder through attacks on cryptographic protocols. In *Proceedings of the 44th ACM Southeast Conference (ACMSE2006)* (pp. 667-672). 10.1145/1185448.1185594

Park, S., Park, H., Won, Y., Lee, J., KISA & Kent, S. (2009). Network working group. BBN Technologies. Retrieved from https://tools.ietf.org/html/rfc5636

Siddiqui, Z., Abdullah, A. H., Khan, M. K., & Alghamdi, A. S. (2013). Smart environment as a service: Three factor cloud based user authentication for telecare medical information system. *Journal of Medical Systems*, *38*(1), 1–14. PMID:24346931

Stinson, D. R. (2006). *Cryptography-Theory and Practice* (3rd ed.). USA: Chapman & Hall/CRC.

Wazid, M., Zeadally, S., Das, A. K., & Odelu, V. (2016). Analysis of Security Protocols for Mobile Healthcare. *Journal of Medical Systems*, *40*(11), 229. doi:10.100710916-016-0596-0 PMID:27640159

This research was previously published in the International Journal of Web Services Research (IJWSR), 15(3); edited by Liang-Jie Zhang; pages 61-81; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Box 1. Code of SRPF in SPDL with results

```
/* SRPF (Secure Robust and Privacy enhanced mobile healthcare Framework).
/* This Mobile Healthcare Framework is based on Wireless PKI, UICC, TPM */
/* Secure Mobile Healthcare Protocol */
/* HC means Mobile Healthcare */
const pk: Function;
secret sk: Function;
inversekeys (pk,sk);
usertype Timestamp;
usertype ACK;
usertype PID,UN,PW,DID,Tp,Td,np,nm,Thc, nhc;
// Protocol description
protocol SRPF(HC,D,P)
{
role HC
{
const nhc: Nonce;
var np: Nonce;
var nd: Nonce;
const Khcp: SessionKey;
const Khcd: SessionKey;
/* Authentication and Key Agreement Protocol */
/* Khcp is the symmetric shared key between patient & HC */
send_1 (HC,D, {UN,PW,PID,Khcp, Thc, nhc}Khcd);
read_2 (D,HC, {ACK,PID,Khcp,Td, nd, Thc, nhc}Khcd);
send_3 (HC,P, {PID,DID,Khcp, Thc, nhc}Khcp);
read_4 (P,HC, {ACK,PID,Khcp,Tp, np, Thc, nhc}Khcp);
claim_HC1 (HC, Secret, Khcd);
claim_HC2 (HC, Secret, nhc);
claim_HC3 (HC, Secret, np);
claim_HC4 (HC, Niagree);
claim_HC5 (HC, Nisynch);
}
role D
{
const nd: Nonce;
var nhc: Nonce;
const Khcd: SessionKey;
const Khcp: SessionKey;
read_1 (HC,D, {UN,PW,PID,Khcp, Thc, nhc}Khcd);
send_2 (D,HC, {ACK,PID,Khcp,Td, nd, Thc, nhc}Khcd);
claim_D1 (D, Secret, Khcd);
claim_D2 (D, Secret, nd);
claim_D3 (D, Niagree);
claim_D4 (D, Nisynch);
}
role P
{
const np: Nonce;
var nhc: Nonce;
const Khcp: SessionKey;
read_3 (HC,P, {PID,DID,Khcp, Thc, nhc}Khcp);
send_4 (P,HC, {ACK,PID,Khcp,Tp, np, Thc, nhc}Khcp);
claim_P1 (P, Secret, np);
claim_P2 (P, Niagree);
claim_P3 (P, Nisynch);
}
}
// An untrusted agent, with compromised key
const e: Agent;
untrusted e;
compromised sk(e);
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 25
An Intelligent Multi-Objective Framework of Pervasive Information Computing

Basant Tiwari
Hawassa University, Ethiopia

Vivek Tiwari
IIIT-NR, India

ABSTRACT

This article describes how electronic healthcare has been the key application of pervasive computing innovations to enhance healthcare quality and protect human lives. Specific patients of constant sicknesses and elderly individuals, by and large, may oblige continuous observing of their wellbeing status wherever they are. In this regard, remote patient monitoring technology plays the various important role through wearable devices to monitor patient's physiological figures. But, this must ensure some broad issues like, wearability, adaptability, interoperability, integration, security, and network efficiency. This article proposes a data-driven multi-layer architecture for pervasively remote patient monitoring that incorporates aforesaid issues. It enables the patient's care at the real time and supports anywhere and anytime requirement with using network infrastructure efficiently.

1. INTRODUCTION

Nowadays, the population of aging people who have incessant illnesses such as diabetes, heart disease and much more is increasing. In addition, youngsters are also experiencing diseases like asthma and obesity due to a lavish lifestyle (Jensen, Jensen, and Brunak, 2012). With all such issues, there is an increase in healthcare costs that demand cost-effective, reliable and pervasive patient healthcare systems. If a patient's health could regularly be observed over an extended amount of time, the doctor could diagnose serious health problems at an early stage and may also provide more accurate treatment.

DOI: 10.4018/978-1-7998-1204-3.ch025

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The Republic of India; a young nation, with the world's most populous community, has battled for making emergency medical care safe and economical (Srinivisan, 2012). Nowadays, Government is thinking about its socioeconomic expansion through Digital India and Skilled India, but it also plagues many challenges (Srinivisan, 2012; Fernald, Capriotti, Daneshjou, Karczewski, and Altman, 2011). The private medical sector of India is holding more or less 81% medical services and 88% medical professionals (Srinivisan, 2012). Health Workforce in India remains inadequate and underutilized, and India started its health reforms journey in the most recent decade. This mission now needs to pick up momentum. Home-based self-monitoring devices are required to become self-reliant for non-invasive and periodic self-monitoring of chronic diseases (Gudwani, Puri, and Vaidya, 2012; Anshari, Alas, and Guan, 2016).

Pervasive healthcare frameworks utilize ubiquitous computing technologies, for example, wearable biomedical sensors with the remote system, to build the modalities and spatiotemporal measurements in which medicinal services administrations can be accommodated enhancing the patient well-being. The most accurate information about the patient at the right time is essential to obtain the best possible service for the patient (Shumer, Tang, Shrivastava, Lombrozo, and Druker, 2008). The goal of Pervasive Healthcare (PH) is to utilize ubiquitous computing technologies to provide continuous health care services from anywhere and outside the scope of traditional medical places, such as hospitals and medical clinics (Kuhn and Prettner, 2016). Pervasive healthcare is aiming to change this conventional way (Visiting a physician /specialist) into pervasive patient monitoring from anywhere and at any time. It exploits the sensing and communication technologies to monitor patient round the clock.

Pervasive healthcare encourages patient-doctor collaboration and can give exact, precise, and upright care to all (Mukherjee, Kolui, and Datta, 2014; Wu et al., 2017; Hart, 2013). This is especially helpful nowadays, since the population is expanding quickly, medical institutions are confronting deficiencies of medical staff, and the cost of healthcare is soaring (Fernald et al., 2011; Hajheydari, Khakbaz, and Farhadi, 2013). The electronic healthcare framework has been the primary application of pervasive computing to enhance healthcare quality and preserve human lives. A number of frameworks consisting of many devices have been invented for remote observation to gain the patient well-being (He, Zeadally, Kumar, and Lee, 2016; Yu and Tseng, 2007; Yuan and Herbert, 2011; Andreu-Perez, Poon,Merrifeld, Wong, and Yang, 2015).

Due to technological advancements in different fields like a wireless network, microchip, integration, and tininess of devices, sensors, and the internet, it is easy to modernize the way of healthcare services and focus on early diagnosis of chronic diseases (Chondrogiannis, Andronikou, Tagaris, Karanastasis, and Varvarigou, 2017).

These systems are structured in such a way that it can react to crisis and manage illness rather than wellness (Banaee, Ahmed, and Loutfi, 2013). In this view, Body Sensor Network (BSN) is a feasible solution, because it is an appropriate blend of modest wearable gadgets connected to the patient's body for observing the patient's physiological information. Sensors continuously screen and gather the patient's information and send it to a remote server through a network that is referred as Database Server (DBS) (Tiwari and Kumar, 2012).

2. ISSUES AND CHALLENGES

A noteworthy challenge is to give continuous and persevering medical services via wearable human services gadgets. Furthermore, numerous patients are more concern about privacy especially, when

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

they realize that their data is open over remote channels. Thus, one most crucial issue is that medicinal services suppliers must secure the patient's data. Significant challenges with Patient Monitoring and Remote Medical Care System are:

1. Since the patient's physiological data are highly sensitive, and BSN is very receptive to attacks; therefore, it must be ensured that the patient's identity should not be exposed and altered as well as the patient's data should not fall into the hands of unauthorized users.
2. We frequently come across data values which are successively similar in content and context, and it does not contribute towards any diagnosis. So, the overall performance of the information computing can be improved by identifying and removing such repeating data (Tiwari and Thakuar, 2016).
3. Electronic health records (EHR) is a fast way to transfer patient data among experts to enhance the quality of clinical issues (Mirani and Harpalani, 2013). For the protection of EHR, Many countries are implementing privacy laws that ensure the confidentiality of such HER data. Furthermore, they even facilitate patient to control and manages the access right of their own records. A data-driven Role Based Access Control (RBAC) might be one of the solutions for data authorization (Sohr, Mustafa, Bao, and Ahn, 2008).
4. Real-time physiological value acquisition.
5. Real-time data analysis and decision making.
6. Strong Decision-Making Policy for patient- authentication system.

This paper incorporates these issues in an architecture which provides solutions for pervasively secured remote patient monitoring. In this regard, we take up an investigation for the development of secure and pervasive architecture which focuses on issues such as minimizing redundant data, compression, and reliable transmission of BSN data with efficient access control policy for accessing the patient's data from the database server. A multi-layered architecture is proposed for pervasively remote patient monitoring to accommodate aforesaid challenges.

3. LITERATURE REVIEW

Majorly, remote patient monitoring systems include a traditional communication system like a telephone network to transmit physiological signals using high-frequency carrier signal (Weiskopf, Hripcsak, Swaminathan, and Weng, 2013; Charles, Gabriel, and Furukawa, 2013; Lau et al., 2011). In spite of the fact that these frameworks permit constant checking, they have an inadequate bandwidth, high telephone charges and degradation due to noise. In recent years, a lot of work has been done in this field.

Yu & Tseng (2007) propose wireless patient monitoring system architecture for a mobile patient monitoring, which provides remote access capability. It consists hardware and software module for the collection and transmission of patient data. They used IEEE standard 802.11 and WMTS for communication purpose and implemented architecture which highly emphasized only on the communication system.

Trinugroho et al. (2012) presented a framework for a real-time web-based health monitoring architecture by using Enterprise Service Bus (ESB). This architecture consists a web-based Service Oriented Architecture (SOA) with ESB that uses subscribe message exchange pattern functionality. PDA like smartphone device is used to implement applications for remote patient monitoring. A wireless body

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

area network has been organized for the accumulation, preparing and transmission of information to various users of patient data including relatives, nurses, and doctors.

Ogunduyile et al. (2013) presented a "Service Oriented Wireless Body Area Networks" (SOWBAN) to monitor physiological data such as SPO2, HR of patients along with the location and activity. Moreover, it contained 3-axis accelerometer and a wireless pulse oximeter (SpO2) as a significant component. Authors discussed the overall functioning of the architecture in three significant layers: (i) Biomedical Sensor (BS) Arduino Layer, (ii) Central Intelligent Node (CIN) layer and (iii) Medical Health Server (MHS) layer. BS Arduino layer implements various body sensors installed on patient's body to collect data. CIN layer receives and process the data and sends it to MHS layer. Where, the data are maintained as a central repository for future reference.

Kang et al. (2013) developed a Wireless Patient Monitoring System (WIPAMS) that monitors the patient's data in real time and transfers to a remote server for further processing. The authors have developed an embedded display board where Android OS was installed. The primary three modules are developed, where, the first module consists sensors, and the second module implements an Android based control logic. The last module poses a dashboard that visualizes collected data. Android OS maintains various threads to monitor vital sign and communication process. Basically, threads are developed for pre-processing of patient's data. Authors conclude their work by providing a specification of developed WIPAMS board (Odroid-X).

Tagad & Matte (Labarère, Bertrand, and Fine, 2014) discussed a Point-of-Care (POC) patient monitoring system that monitors out-patients who require care. They enhanced the traditional system with automatic remote monitoring. A fundamental wireless communication system is developed to transfer data into a centralized data repository. These data are generated by wearable sensors, Bluetooth and Zigbee. It also provides an alarm generation system, which informs to concern doctor if things go wrong via SMS.

Mukherjee et al. (2014) presented an architecture for electronic health care which uses wireless infrastructure with the smartphone-like device for collection, monitoring, and analysis of patient's data at a real-time emergency situation. The architecture consists of three layers; (i) perception layer, (ii) Middleware & APIs layer and (iii) E-Health application & service layer. Perception layer consists a set of medical sensors to collect the patient's physiological data. It measures the information like id, value measuring unit, timestamp, etc. and convert this data into metadata and finally send it to the next layer. Middleware & APIs layer implements cloud infrastructure with a set of APIs for (i) patient registration, (ii) profile creation, (iii) storage, (iv) patient history queries, etc. E-Health Application & Service Layer provide outsourcing services for patient's data and healthcare advice services to the patient. Authors also describe various security requirements for patient's medical data.

Yuan & Herbert (2011) introduced a web-based Context-Aware Real-time Assistant (CARA) architecture to monitor patient's physiological data with contextual information. Authors demonstrated 3-tier architecture for real-time patient monitoring in which body area network, the Internet and CARA server are implemented on each layer respectively. Protocol of CARA system is developed through XML. Security and privacy issues for protecting patient's privacy are also included in the work.

Various architectures for remote patient monitoring are introduced, but many significant issues are not considered effective as a whole package such as access control, redundancy reduction and security of patient's data as well as network issues like better utilization of network bandwidth.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. MULTI-OBJECTIVE FRAMEWORK

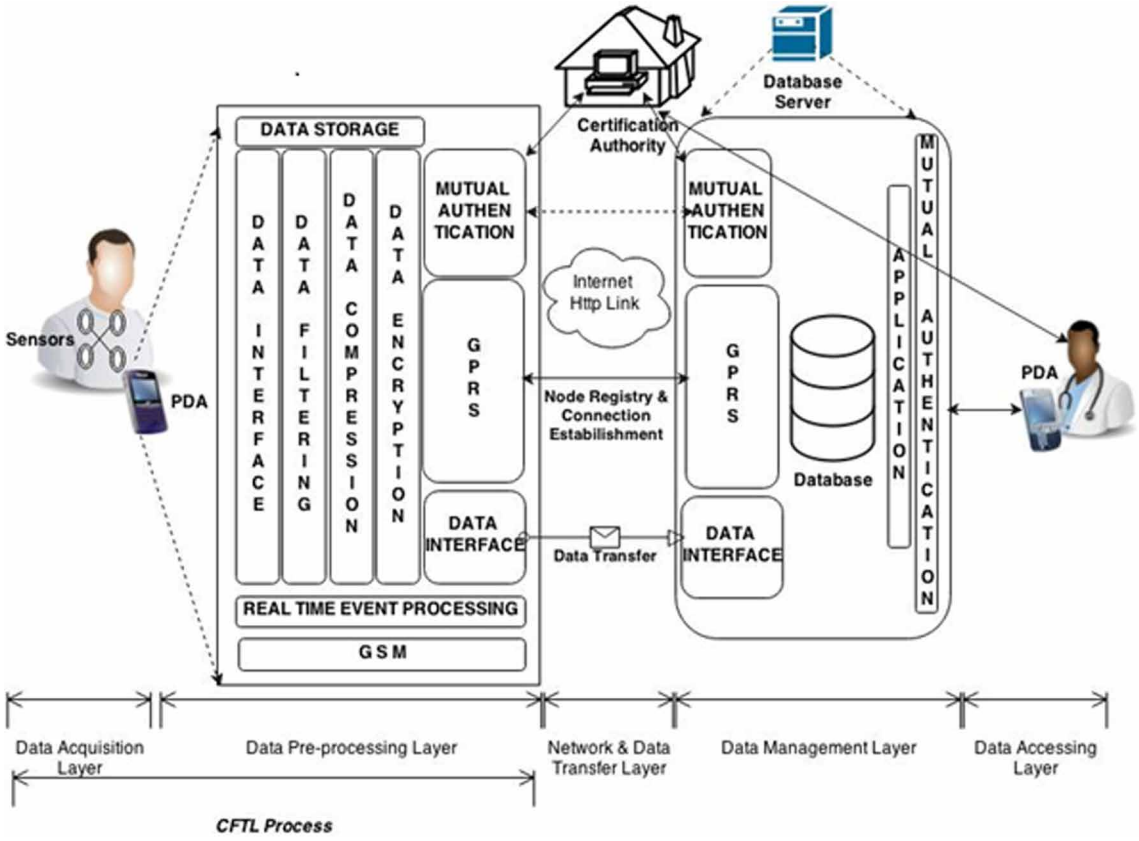
This section is devoted to the layered architecture for pervasive healthcare for remote patient monitoring and is presented in Figure 1.

Majorly, it comprises of five layers as:

1. Data Acquisition Layer, (Layer 0)
2. Data Pre-processing Layer, (Layer 1)
3. Network and Data Transfer Layer, (Layer 2)
4. Data Management Layer, (Layer 3) and
5. Data Accessing Layer (Layer 4).

Data Acquisition and Data Pre-processing (Layer 0 and 1) work at the patient site and collectively launch a process called CFTL process (Collect, Fuse, Transform and Load process). Data Management Layer (Layer 3) works at the remote Database Server (DBS) where all patient data are stored. Data Access Layer (Layer 4) work at the healthcare provider site and Network and Data Transfer Layer (Layer 2) use network infrastructure to transfer the data from patient site to DBS.

Figure 1. Patient remote monitoring system architecture



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CFTL process includes sensor devices installed in the patient's body and the personal server called PDA (Personal Device Assistant). In CFTL process, sensors collect the patient data with fixed interval according to their functionality and specifications. Received data is pre-processed at PDA, where different types of patient's data are filtered, compressed, encrypted, fused and transformed into a single packet and is loaded as a single entity. The primary purpose of CFTL Process are:

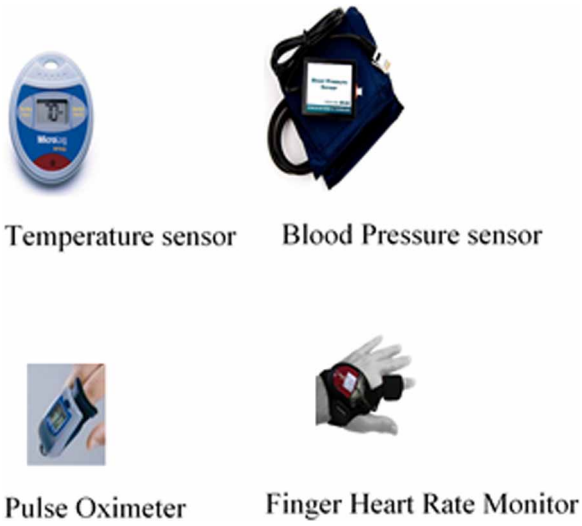
1. Increase the utilization of the bandwidth of the wireless channel
2. To reduce redundant values transmission
3. Easy to maintain authenticity, confidentiality, and integrity of patient's data
4. Energy and cost saving

4.1. Data Acquisition Layer (Layer-0)

It is primarily used for sensing and collecting the data and is equipped with the wearable body Sensors as depicted in Figure 2. These sensors measure various physiological and environmental signals. Sensor devices communicate wirelessly with network gateways (Smart Phones or PDAs) which store, analyse and communicate vital information to Layer-1 in real time.

Such sensor cannot work by themselves; they need a microprocessor, local storage and a wireless radio for functioning. With these things, the sensor becomes a sensor node which is an embedded device. There are various mote platforms available, each having a different configuration like processor and radio specification, but device configuration and specifications are more or less similar. TinyOS is an embedded operating system specially developed for sensors and do not have full functionality of a regular operating system (Linderman, Mechitov, and Spencer, 2013). The nodes which can be used are the Moteiv-Tmote-Sky based on the TelosB platform. The Tmote Sky has a 8MHz TI MSP430 microprocessor which has 10kB of RAM and 48kB of flash memory available. This microcontroller is

Figure 2. Various Sensors



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

used in wireless sensor network devices. Wireless communication takes place among nodes using the Chipcon CC2420 radio. The CC2420 is a 250kbps, IEEE 802.15.4 (Zigbee) compliant radio operating at 2.4GHz. These motes run the TinyOS operating system (Linderman et al., 2013).

Data acquisition from the sensor is possible by taking the set of sensors S_k, where k is the number of the sensor implanted over patient's body such as $S_k = \{s_1, s_2, ..., s_k\}$. Each sensor senses the physiological value as per their functionality and specifications, i.e. Skpv. At last, data is sent to the next layer for further processing.

4.2. Data Pre-Processing Layer, (Layer-1)

This layer is a part of CFTL process and installed on PDA for data pre-processing. PDA is a processing and connectivity device that can interact with the sensor nodes through various IEEE standards like Bluetooth and Zigbee at one end and another end. It may also contain a user interface or an interface to a connecting network. PDA provides more processing power than the sensor node and uses its internal memory to store the data and program to execute. PDA memory can be expanded with a flash memory card for real-time storage and pre-processing of sensor data.

In the proposed architecture, this layer consists of various sub-layers for patient data pre-processing such as:

1. Data Interface Sub-layer
2. Data Filtering Sub-layer
3. Data Compression Sub-layer
4. Data Encryption Sub-layer

Collected patient data are passed through these sub-layers and processed according to the functionality of each sub-layer. The functionalities of each sublayer are:

4.2.1. Data Interface Sub-Layer

This is an interface between sensor devices at layer-0 and other sub-layers of layer-1. It collects the various sensed data and sends to data filtering sublayer. PDA provides multiple wireless interfaces that can be used for healthcare application. Cellular and Wi-Fi, Bluetooth, NFC are wireless communication protocols included in PDA, where Cellular and Wi-Fi are used for Internet connectivity (Jantunen, 2012). Bluetooth and NFC can be used connecting PDA to the sensor devices. Wi-Fi is expensive due to the high data rate and high-power consumption. It also requires continuous energy sources at both ends (at PDA and Sensor devices) for the communication (Chang, 2014). In this way, Bluetooth can be a reasonable choice made up especially for sensor devices. Another feasible option for data communication between the sensor device and PDA may be Zigbee protocol and specially intended for low information rates and low power consumption (Wang, Chen, Lin, Fang, and Lee; 2015). Zigbee technology is simple and less expensive than Wi-Fi and Bluetooth. Nowadays, Zigbee based PDA are available in the market but does not have vast expansion as with Bluetooth, i.e., it is freely available on all smartphones.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4.2.2. Data Filtering Sub-layer

It has been observed that at some point in time non-rehashed, yet repetitive values transmitted unnecessarily. These repetitive values do not give to any conclusions on the ground, i.e., it demonstrates the same state of the patient for a broadened timeframe. These values ultimately increase overhead on preprocessing, communication time and database server memory.

In this view, data filtering technique reduces the communication cost, especially in the case of energy constrained devices. Initially, screening process classifies the data according to sensors values, i.e., for example, Heart-Rate sensor's physiological values are grouped; similarly, Blood-Pressor sensor's physiological values are arranged, etc. Furthermore, Standard Deviation (SD) method is applied to filter the physiological data (Tiwari and Kumar, 2015). Here, we have considered 'Window Size' concept for the range of physiological values, such as heart rate will be 60-80 bpm for a normal adult male is defined as normal, blood pressure 80-120 mm/gm, etc. Such values are standard and having a proper interpretation in the medical field. This concludes that the span of physiological values is fixed, and hence the window size will also be fixed (i.e. the elasticity of window is zero). The proposed method enlarges the elasticity of window at both ends (MIN/MAX) without upsetting the performance. This improves data reduction automatically. An elasticity of the window is calculated by representing the patient's HR data as vector (h_1, h_2, h_3, …, h_n) at the PDA over a particular span of time and then calculate the mean over that data $h' = \sum_{I=1}^{n} h_i / n$. The deviation is measured for each sensed value by subtracting the mean value i.e. $D_I = h_i - h'$ for each h_i and variance is calculated as, i.e., $\left(\sum_{I=1}^{n} D_i^2\right) \Big/ n$. SD is calculated by taking the square root of variance, i.e. $SD = \sqrt{(\sum_{I=1}^{n} D_i^2) / n}$. This SD is considered as an elasticity at both upper and lower end, and so, it enlarges the elasticity of HR value. It has been shown that the amount of data transmitted is reduced up to 40%.

4.2.3. Data Compression Sub-Layer

Data compression techniques can save energy as well as provide efficient utilization of network bandwidth by enormously diminishing the size of information, sent on the remote channel. Hence, there requires a lossless compression technique which can reduce the data in significant size without losing their meaning or any byte especially, when dealing with patient's physiological data. Aggregated Deflate RLE (ADR) is found the best suitable algorithm to this end (Tiwari et al., 2012). ADR offers to improve compression ratio and lowest compression time concerning different standard method such as Huffman, LZW, and Deflate. ADR implements look-ahead buffer of LZ77, that is an ingredient of Deflate. It works well with small file sizes (1KB, 2KB, 5KB, 10 KB and 20 KB) and hence, suitable for BSN data. ADR is tested on many parameters such as Compression and Decompression Time (CT/DT), Compression Ratio (CR), Factor (CF) and Saving Percentage (SP). Intel core i3 IInd Generation processor at 2.30 GHz with 32-bit OS and 2 GB RAM is taken for implementation of this approach. Various observation is taken which conclude that ADR algorithm is found to be the best out of other algorithms taken into consideration. It has been found that ADR provides high CF and SP with lowest CR. These are main aspects when

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

dealing with data compression. It is also noticed that CT and DT are found comparatively more, but it is in the millisecond that is acceptable. The beauty of the proposed system is that data compression is applied to less sensitive information to keep the trust on analytical outcomes. It is best suitable and efficient because of the contextual data partition. An administrator can decide the contexts (less sensitive data) based on domain knowledge where data compression is required or not.

4.2.4. Data Encryption Sub-Layer

As the name suggests, this sub-layer encrypts the compressed data which is pushed by the just preceding data compression sublayer. Data compression layer ensures information security of physiological data which flow through a network. There are various symmetric, and asymmetric encryption algorithms exist as a choice, but the algorithm with the minimum functional requirement is suggested to choose. In this view, a Physiological Value based Encryption and Mutual Authentication (PVEMA) mechanism is best suitable that facilitates mutual authentication as well as data encryption (Tiwari and Kumar, 2013). PVEMA is based on the Elliptic Curve Cryptography (ECC) with Message Authentication Code (MAC) and Symmetric Encryption technique for ensuring integrity, confidentiality and authenticity (i.e., previously stored Physiological value is used as input for key generation.)

PDA executes ECIES algorithm which is a variation of ECC that uses generated keys for encryption. In the process of PVEMA, we have found domain parameters a, b, G and n using chosen elliptic curve then key generation process executes which leads to Encryption/Decryption process. The mutual authentication process initially calls a Certification Authority when the system is setup that is described in section-5 of this paper. The encryption process takes data M and shared physiological values S1 and S2 and a triplet (R, C, D) is generated as output or encrypted message. The process initiated by taking a random number k i.e. $k \in [1, n- 1]$ and then it calculates the $R = kG$. Here a Key Derivation Function (KDF) is used to derive key for symmetric encryption of message which generated C and then MAC is used for tagging operation which produces the D. These three values are grouped into triplet as an output of encryption procedure. Now, this triplet is sent to Database Server (DBS) where physiological values are stored for future use. DBS executes decryption process using same ECIES algorithm. This uses the same generated keys and parameters that are shared during system setup and key deployment procedure. Here the triplet (R, C, D) is taken as input with shared physiological values S1 and S2 and output message M is generated. Here the same KDF is used for key generation and same MAC for tag verification. Once tag verified as matched, the symmetric decryption process generates the message M. This work ensures data confidentiality, integrity, and unforgeability. It also avoids various attacks like reply attack, Man in the Middle attack, chosen cipher attack. Now the patient data at DBS is stored in the form of EHR.

4.3. Network and Data Transfer Layer (Layer-2)

Layer-1 combines all the data after pre-processing as a single unit in the form of packets. This packet contains all patient's EHR with the address of DBS. Now PDA transmits it to DBS. The pervasive nature of the Internet enables patient's data to be transmitted in real time. In the proposed architecture, layer-2 is responsible for transmission. This layer implements a client-server architecture, where PDA works as a client while DBS works as a server. The transmission uses TCP/IP protocol suite for transmission of patient's data. This layer uses General Packet Radio Service (GPRS services) to make connections

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

between PDA and DBS and then to transfer of data also. PDA uses 3G or 4G cellular services with the data rate of minimum 2 Mbit/s. GPRS enabled PDA user SIM card slot and allowed transfer of files to another device like DBS.

4.4. Data Management Layer (Layer-3)

Vital Sign sensed and collected by a sensor at layer-0, after pre-processing at layer-1, is forwarded to a remote and centralized back-end repository that is implemented at layer-3 called data management layer which aggregated all patient information called Healthcare Records (EHR) in the form of centralized database. A Healthcare record comprises of Personal information, Insurance Information, and Medical data. The personal information contains patient-id, patient-name, gender, age, address, etc. Insurance information includes insurance date, insurance type, amount, claimed amount, etc., while Medical data comprise of patient's collected sensor data, his/her lab conducted data and also Sensitive data like HIV, mental and sexual data. This heterogeneity in data is at context and format level and so, proposed to handle by dividing the data as per underlying contexts. This repository works as a Database Server (DBS) or server repository. This remote and centralized database server manages all information in a real-time manner. It stores and shares the health-related data of a patient. This server repository works only online and provides access to all patient data to only authorized users like health care providers in an emergency situation for their clinical trials, insurers, relatives and some time to research students.

Database server interface implements various software system calls that are used by the users or administrator to manage the database. System calls may be inserted (), querying (), update () etc. All the patient information is accessed only after establishing secure connection and authentication of the user.

4.5. Data Accessing Layer (Layer-4)

This layer deals with accessing and transferring medical data from a remote database server on the request of EHR users. This layer provides the convenient methods access patient information among various users while keeping the access verifiability (Sohr et al., 2008. Because such medical history data pose sensitive information and it forced to implement methods to ensure confidentiality. So this layer performs various approaches to access control. These access control approaches may be Discretionary Access Control (DAC), Mandatory Access Control (MAC) or Role Based Access Control (RBAC). This layer combines works with the CA to prevent unauthorized access and to provide security and privacy to EHRs.

In continuation, "Role-Based Access Control through On-Demand Classification of Electronic Health Record" (Tiwari and Kumar, 2015) is based on Access Control through Classification (AC^2) algorithm that uses the logistic regression to classify the patient's data as per underlying user's role. Furthermore, various roles are used as the class, and each class has their feature vector for classification of HER, and this describes the user authority. EHR is categorized majorly three access level (Personal data, Insurance, and EHR). Multi-level data categorization is also possible as far as EHR is concerned, for example, personal information may be considered at first level that may further be categorized into the name, date of birth, gender, age, address and live status, etc. Furthermore, binomial feature 'live status' (dead, alive) containing sensitive information such as if death is true to have further one more level granularity like death reason, location and considered to access to the personal doctor. Similarly, Nurse may have access to personal patient data and sensed physiological values, the Normal user has right to access merely abstract personal information and so on. The objective of proposed work was to describe user's au-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

thorities only. AC2 algorithm takes User role (Ur'), Patient-Id (Pid), Global Feature Vector (Gfv), Access Policies (AP) and class label (C) as input and it generates Class (Dt'). Where, Dt' $\in$ Dt as output where Dt is a given patient's record and Dt' is subset of Dt with respect of given user's role. Initially, the feature-vector is set to Null and the according to user role the selection is performed on global feature vector (Gfv) (i.e. Fv = σ user role= Ur (Gfv)) and selected features are stored at the Fv. Now from global EHRDB, the patient's EHRDB is created using patient-id and selection operator over EHRDB i.e. PEHRDB = σ Patient-id= Pid (EHRDB). Now 1:1 mapping of user role to assigned policy is defined (g: Ur $\rightarrow$ Pu) and then m-dimensional feature vector space is created. Now as per the user policy attribute values are fetched from the Patient's EHRDB by Scanning [Dt] with respect to Vur where Vur is the set of values of respective features. The bottom-up paradigm has been adopted to better fitment for access control where data values itself comes together to make classes. It also provides automatic access control model whereas and when user login the system, according to user's role the fraction of global data is grouped. The work also provides on-demand data classification whereas and when user role change the data classification is modified since the user's access rights will either increase or decrease.

5. CERTIFICATION AUTHORITY

Nowadays, the rapid growth of internet transactions is emphasizing the importance of secure authentication of network entities. For reliable and authorized data transmission between PDA & database server and healthcare providers & database, the involvement of trusted third party, such as a Certification Authority (CA) is necessary. CA is responsible for providing essential primitives for mutual authentication as well as secure transmission of patient's data among the EHR users. CA works online in real-time that issues and manages security credentials like security keys for maintaining the integrity and authorized access. The certification authority is governed and controlled by government bodies which provide trustworthy relations between network entities. CA contains the identity of network entities like domain owner and online servers for authenticated transactions.

In this regard, similar approach PVEMA (Tiwari et al., 2013) has been taken, where, the moment patient is equipped with a new BSN, it establishes a connection with DBS and when a link is set up, both PDA and DBS get in touch with CA to receive a certificate. For certificate creation, PDA and DBS both shares their public keys (Q_{pda} and Q_{dbs}) with CA. Now, CA generates the signature using their private key. CA used Elliptic Curve Digital Signature Algorithm to generate certificates and sends certificates (r_{pda},s_{pda}) for PDA and ($_{rdbs}$,s_{dbs}) for DBS with own public key Q_{ca}. When both the parties start mutual authentication process between them, these certificates and CA's public key is used.

6. CONCLUSION

We have investigated and developed a safe and pervasive model for patient monitoring and care. Here all requirements regarding patient privacy, security of data, removing redundancy without loss of data semantics, energy and bandwidth saving by data compression, reliable access control mechanism, etc. are fulfilled. In conclusion, presented work attempts to maintain privacy, secrecy, authenticity, and integrity of patient's data with energy and bandwidth saving which gives a significant cost reduction for pervasive patient monitoring. The presented multi-objective framework seems as an analytical system

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

for Healthcare for budgetary, operational, quality and process enhancement and works as the source framework for populating different contextual data. It is a healthcare analytics tool that dispenses with the need to compose SQL, empowers publishing of reports and fit for creating an extensive range of visualization (charts, crosstabs, tables, and graph).

REFERENCES

Andreu-Perez, J., Poon, C. C., Merrifield, R. D., Wong, S. T., & Yang, G. Z. (2015). Big data for health. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1193–1208. doi:10.1109/JBHI.2015.2450362 PMID:26173222

Anshari, M., Alas, Y., & Guan, L. S. (2016). Developing online learning resources: Big data, social networks, and cloud computing to support pervasive knowledge. *Education and Information Technologies*, *21*(6), 1663–1677. doi:10.100710639-015-9407-3

Banaee, H., Ahmed, M. U., & Loutfi, A. (2013). Data mining for wearable sensors in health monitoring systems: A review of recent trends and challenges. *Sensors (Basel)*, *13*(12), 17472–17500. doi:10.3390131217472 PMID:24351646

Chang, K. H. (2014). Bluetooth: A viable solution for IoT? *IEEE Wireless Communications*, *21*(6), 6–7. doi:10.1109/MWC.2014.7000963

Charles, D., Gabriel, M., & Furukawa, M. F. (2013). Adoption of electronic health record systems among US non-federal acute care hospitals: 2008–2013. *ONC data brief*, *9*.

Chondrogiannis, E., Andronikou, V., Tagaris, A., Karanastasis, E., Varvarigou, T., & Tsuji, M. (2017). A novel semantic representation for eligibility criteria in clinical trials. *Journal of Biomedical Informatics*, *69*, 10–23. doi:10.1016/j.jbi.2017.03.013 PMID:28336477

Fernald, G. H., Capriotti, E., Daneshjou, R., Karczewski, K. J., & Altman, R. B. (2011). Bioinformatics challenges for personalized medicine. *Bioinformatics (Oxford, England)*, *27*(13), 1741–1748. doi:10.1093/bioinformatics/btr295 PMID:21596790

Gudwani, A., Mitra, P., Puri, A., & Vaidya, M. (2012). *India healthcare: inspiring possibilities, challenging journey*. New York, NY: McKinsey & Company.

Hajiheydari, N., Khakbaz, S. B., & Farhadi, H. (2013). Proposing a Business Model in Healthcare Industry: E-Diagnosis. *International Journal of Healthcare Information Systems and Informatics*, *8*(2), 41–57. doi:10.4018/jhisi.2013040104

Hart, V. (2013). Hospital IT Sophistication Profiles and Patient Safety Outcomes: A Comparison of Three States. *International Journal of Healthcare Information Systems and Informatics*, *8*(1), 17–36. doi:10.4018/jhisi.2013010102

He, D., Zeadally, S., Kumar, N., & Lee, J. H. (2016). Anonymous authentication for wireless body area networks with provable security. *IEEE Systems Journal*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jantunen, I. (2012). *System Architecture for Mobile-phone-centric Ambient Intelligence Applications* [Doctoral Dissertation].

Jarecki, S., & Liu, X. (2009). Private mutual authentication and conditional oblivious transfer. In *Advances in Cryptology-CRYPTO 2009* (pp. 90–107). Springer Berlin Heidelberg. doi:10.1007/978-3-642-03356-8_6

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: Towards better research applications and clinical care. *Nature Reviews. Genetics*, *13*(6), 395–405. doi:10.1038/nrg3208 PMID:22549152

Kang, J., Yoo, S., & Oh, D. (2013). Development of a Portable Embedded Patient Monitoring System. *International Journal of Multimedia & Ubiquitous Engineering*, *8*(6), 141–150. doi:10.14257/ijmue.2013.8.6.14

Kuhn, M., & Prettner, K. (2016). Growth and welfare effects of health care in knowledge-based economies. *Journal of Health Economics*, *46*, 100–119. doi:10.1016/j.jhealeco.2016.01.009 PMID:26918295

Labarère, J., Bertrand, R., & Fine, M. J. (2014). How to derive and validate clinical prediction models for use in intensive care medicine. *Intensive Care Medicine*, *40*(4), 513–527. doi:10.100700134-014-3227-6 PMID:24570265

Lau, E. C., Mowat, F. S., Kelsh, M. A., Legg, J. C., Engel-Nitz, N. M., Watson, H. N., & Whyte, J. L. (2011). Use of electronic medical records (EMR) for oncology outcomes research: Assessing the comparability of EMR information to patient registry and health claims data. *Clinical Epidemiology*, *3*, 259–272. PMID:22135501

Linderman, L. E., Mechitov, K. A., & Spencer, B. F. Jr. (2013). TinyOS-based real-time wireless data acquisition framework for structural health monitoring and control. *Structural Control and Health Monitoring*, *20*(6), 1007–1020. doi:10.1002tc.1514

Mirani, R., & Harpalani, A. (2013). Business benefits or incentive maximization? impacts of the medicare EHR incentive program at acute care hospitals. *ACM Transactions on Management Information Systems*, *4*(4), 20. doi:10.1145/2543900

Mukherjee, S., Dolui, K., & Datta, S. K. (2014). Patient health management system using e-health monitoring architecture. In *IEEE International on Advance Computing Conference (IACC)* (pp. 400-405). IEEE. 10.1109/IAdCC.2014.6779357

Ogunduyile, O. O., Zuva, K., Randle, O. A., & Zuva, T. (2013). Ubiquitous healthcare monitoring system using integrated triaxial accelerometer, spo2 and location sensors. *International Journal of UbiComp*, *4*(2), 1–13. doi:10.5121/iju.2013.4201

Shumer, D., Tang, N., Shrivastava, S., Lombrozo, P., & Druker, J. (2008). U.S. Patent No. D578, 645. Washington, DC.

Sohr, K., Mustafa, T., Bao, X., & Ahn, G. J. (2008). Enforcing role-based access control policies in web services with UML and OCL. In ACSAC 2008 Computer Security Applications Conference (pp. 257-266). IEEE. 10.1109/ACSAC.2008.35

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Srinivisan, R. (2012). Health Care In India - Vision 2020, Issues And Prospects, Planning commission. Govt of India. Retrieved Feb. 2016 from http://planningcommission.gov.in/reports/genrep/bkpap2020/26_bg2020.pdf

Tiwari, B., & Kumar, A. (2012, September). Aggregated Deflate-RLE compression technique for body sensor network. In *2012 CSI Sixth International Conference on Software Engineering*.

Tiwari, B., & Kumar, A. (2013). Physiological Value Based Privacy Preservation of Patient's data Using Elliptic Curve Cryptography. *Health Informatics–An International Journal*, *2*(1), 1–14. doi:10.5121/hiij.2013.2101

Tiwari, B., & Kumar, A. (2015). Role-based access control through on-demand classification of electronic health record. *International Journal of Electronic Healthcare*, *8*(1), 9–24. doi:10.1504/IJEH.2015.071637 PMID:26559071

Tiwari, B., & Kumar, A. (2015). Standard deviation (SD)-based data filtering technique for body sensor network data. *International Journal of Data Science*, *1*(2), 189–203. doi:10.1504/IJDS.2015.072420

Tiwari, V., & Thakur, R. S. (2016). Pattern Warehouse: Context Based Modeling and Quality Issues. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, *86*(3), 417-431.

Trinugroho, Y. B. D., Rasta, K., Nguyen, T. H., Fensli, R. W., & Reichert, F. (2012). A Real-Time Web-Based Health Monitoring System Based on Enterprise Service Bus. In *Proceedings of 2012 11th IADIS International Conference on WWW/Internet* (pp. 165-172).

Wang, L. H., Chen, T. Y., Lin, K. H., Fang, Q., & Lee, S. Y. (2015). Implementation of a wireless ECG acquisition SoC for IEEE 802.15. 4 (ZigBee) applications. *IEEE Journal of Biomedical and Health Informatics*, *19*(1), 247–255. doi:10.1109/JBHI.2014.2311232 PMID:25561447

Weiskopf, N. G., Hripcsak, G., Swaminathan, S., & Weng, C. (2013). Defining and measuring completeness of electronic health records for secondary use. *Journal of Biomedical Informatics*, *46*(5), 830–836. doi:10.1016/j.jbi.2013.06.010 PMID:23820016

Wu, P. Y., Cheng, C. W., Kaddi, C. D., Venugopalan, J., Hoffman, R., & Wang, M. D. (2017). –Omic and Electronic Health Record Big Data Analytics for Precision Medicine. *IEEE Transactions on Biomedical Engineering*, *64*(2), 263–273. doi:10.1109/TBME.2016.2573285 PMID:27740470

Yu, H. C., & Tseng, S. M. (2007). A wireless based sensor for patient monitoring system with remote diagnostic. In *Third International Conference on Networking and Services 2007* (pp. 92-92). IEEE. 10.1109/ICNS.2007.24

Yuan, B., & Herbert, J. (2011). Web-based real-time remote monitoring for pervasive healthcare. In *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)* (pp. 625-629). IEEE. 10.1109/PERCOMW.2011.5766964

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 13(4); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 15-27; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 3
Tools and Technologies

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 26
Detecting Human Diseases Relatedness: A Spreading Activation Approach Over Ontologies

Said Fathalla
Bonn University, Germany & Alexandria University, Egypt

ABSTRACT

Due to the ubiquitous availability of the information on the web, there is a great need for a standardized representation of this information. Therefore, developing an efficient algorithm for retrieving information from knowledge graphs is a key challenge for many semantic web applications. This article presents spreading activation over ontology (SAOO) approach in order to detect the relatedness between two human diseases by applying spreading activation algorithm based on bidirectional search technique. The proposed approach detects two diseases relatedness by considering semantic domain knowledge. The methodology of the proposed work is divided into two phases: Semantic Matching and Diseases Relatedness Detection. In semantic matching, diseases within the user-submitted query are semantically identified in the ontology graph. In diseases relatedness detection, the relatedness between the two diseases is detected by using bidirectional-based spreading activation on the ontology graph. The classification of these diseases is provided as well.

1. INTRODUCTION

Usually, computers were used to perform tasks with none sort of intelligence or semantics, thus there is a necessity to make computers as intelligent as human (Antoniou & Harmelen, 2004). The defect of non-semantic applications is painted by the statement "lack of semantics" particularly when the talk is concerning information retrieval. Numerous semantic-based applications require an efficient algorithm for querying large ontologies. The semantic web gives the meaning of knowledge, making software systems can perform intelligent tasks instead of users using a semantic content called ontology (Berners-

DOI: 10.4018/978-1-7998-1204-3.ch026

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lee, Hendler & Lassila, 2001). Ontology is a repository in which information is organized and used in semantic-based applications (Shadbolt, Berners-Lee & Hall, 2006). Therefore, ontologies are widely used as a semantic knowledge base in different areas of computer science such as information retrieval (Oberle, 2014), text classification (Fathalla, Hassan & El-Sayed, 2012), scholarly communication (Fathalla, Vahdati & Auer, 2017) and bioinformatics (Fathalla & Kannot, 2017).

The use of ontologies in the field of health informatics has become a mainstream activity within bioinformatics due to the vast growth of healthcare system. In bioinformatics, ontology is used for the representing and organizing medical vocabularies.

In neurophysiology, communications between neurons are demonstrated by a way of activation, which propagates from one neuron to another via connections called synapses to transmit information using signals (Ali, 2009). This phenomenon observed in the nervous systems of living organisms and called spreading activation. This phenomenon was later exploited in Artificial Intelligence as a method for searching associative networks, neural networks, or semantic networks (Schumacher, et al., 2008). Spreading activation in a generic form is a set of methods suitable for mining multidimensional networks with oriented weighted links. Collins and Loftus (1975) discussed in their research that the Spreading Activation (SA) runs on semantic networks (Anderson et al., 1983) and is used for information retrieval process (Burns et al., 1981). The spreading activation algorithm is suitable to work with incomplete data and with large datasets. It runs on a graph structure that comprises a set of nodes connected by edges. The concepts are nodes which have an activation value and the relations between them are represented by edges between these nodes. An activation value is assigned to each node in the graph and then the algorithm spreads to the nodes with the highest activation value. The algorithm runs in a set of iterations and terminates when a stopping condition is reached. The output of the algorithm is the list of nodes that have been fired according to their activation values. For each iteration or cycle, there are three substantial actions:

- The list of nodes is expanded by adding adjacent nodes of the latest activated nodes.
- The activation value of each node in the list is recomputed based on the activation value of the node itself and the weight of links which exist between other nodes.
- The list is filtered by excluding the nodes with activation values less than a given threshold.

This paper intended to demonstrate how to identify whether two human diseases are related to each other or not. If so, what are the set of diseases connecting them. For instance, is there a relatedness between “Vasculogenic impotence” and “Transvestism”? And if so, what is the path (set of diseases) of the relatedness?

Finding a relatedness between diseases is important because physicians can treat patients not only based on the symptoms they suffer from but treat the real cause of this disease which may be related to another disease that causes these symptoms, i.e., physicians can treat the cause disease, not the symptoms.

To the best of our knowledge, it has not yet been established whether biomedical ontologies could be used to detect relations between human diseases. After reviewing the literature, we believe that this approach is a first such attempt of detecting the relations between diseases using ontologies combined with Descriptive Logic (DL) rules. We have added several Descriptive Logic (DL) rules to the ontology to be able to infer the semantic relations between diseases which are not explicitly provided. In addition, we have also developed the semantic matching and relatedness detector algorithm for identifying user-submitted diseases in the ontology graph and detecting the relatedness.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The remainder of this article is organized as follows. Section 2 gives a brief overview of the related work. Section 3 presents the disease ontology utilized by the proposed approach which is used as a semantic knowledge base. In Section 4, the methodology of the proposed work is discussed. In Section 5, the workflow of the proposed approach is described. In Section 6, a running example is presented. Finally, the conclusion and the directions for future work are presented in Section 7.

2. RELATED WORK

There are a numerous amount of literature on biomedical knowledge management and medical decision making due to the explosion of biomedical knowledge over the last two decades. Therefore, biomedical knowledge available on the web is growing considerably as most of the biomedical research papers are published online. Semantic matching is used to match information which is semantically related based on the meaning of the data rather than keyword-based matching (Giunchiglia et al., 2007). This could be achieved by analyzing the meaning of concepts, not just the labels. There are essentially two approaches to make use of the hierarchical ontology structure for determining the semantic similarity between concepts: edge-based or node-based approach (Bernstein, 2005). In recent years there has been considerable interest in semantic matching (Fathalla & Kannot, 2017; Giunchiglia et al., 2004; Hasan and Curry, 2014; Gonzalo et al., 2014; Guo et al., 2014; Wu et al., 2017). Some examples of individual approaches addressing the matching problem can be found in (Dhamankar et al., 2004; Do & Rahm, 2002). Identifying named text features in user-submitted query is a crucial procedure for efficient semantic matching and text analysis. The analysis may include term normalization; phrase analysis, word sense analysis, and the matching may be performed on phrase aspect and word sense aspect (Kacholia, et al., 2005). In their cutting-edge paper of 2004, Kanavos, Theodoridis, and Tsakalidis (2014) proposed a metasearch engine over PubMed[1]. It provided an accurate clustering of results according to their specific topic using a biomedical ontology, named MeSH[2] in addition to RxNorm[3] that is a US-specific terminology in medicine that semantically links clinical drug vocabularies and supports the interoperation between them. Ngo, Cao and Le (2010) proposed an ontology-based vector space model for semantic annotation and semantic search which combines different ontologies. It took the advantage of ontological features of both named entities and WordNet vocabularies and develops a spreading activation algorithm for query expansion. As anticipated, their experiments demonstrated that the proposed model is better than the solely keyword-based model and also better than the ones using only WordNet or named entities. In addition, they merged various ontologies to improve the semantic search process. De Maio et al. (2011) proposed a project named ODINO which uses a fuzzy knowledge approach for disease diagnosis which supports medical decision making. ODINO has three main features which are a faceted search of diseases through taxonomy constraints, disease catalog browsing, and preliminary medical diagnosis.

Many efforts have been made to represent medical information as a semantic knowledgebase for further processing by semantic applications. Despite the existence of many medical ontologies, it is important to reuse and integrate the ontologies and to establish mappings between their concepts. Recently, there is a growing interest in ontology matching and integration (Shvaiko and Euzenat, 2013; Euzenat, 2011). Shvaiko and Euzenat (2013) surveyed the state of the art of ontology matching and addressed some worthy challenges for ontology matching techniques and also, they conjecture that significant improvements can be obtained. In addition, they analyzed the results of recent ontology matching evaluations. Below some of the famous medical-related ontologies will be listed.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- *Human Disease Ontology* (DO)[4] is a standardized biomedical ontology which contains a considerable number of disease terminologies. Compound disease names, conjunction of two disease terms, are not included.
- *Vaccine Ontology* (VO)[5] contains more than 2000 vaccines and relationships between them (He et al., 2009). VO is written in the Web Ontology Language (OWL).
- *Infectious Disease Ontology* (IDO)[6] is a biomedical ontology which represents infectious diseases and relations between them.
- *Ontology for Biomedical Investigations* (OBI)[7] is an ontology represents life-science and clinical practice concepts.
- *Ontology for General Medical Science* (OGMS)[8] (formerly, Clinical Phenotype Ontology) is an ontology represents diseases and symptoms for human diseases.

3. THE DISEASE ONTOLOGY

Resource Description Framework (RDF) data model is used to describe data semantically, including concepts and relationships between them. The main purpose of developing ontology is to use it as a semantic knowledge base representation for identifying concepts in a specific domain and find out its semantics and the relationships between them (Fathalla, Hassan & Kholief, 2013). The authors of Köhler (2014) and Croft (2014) discuss in their research that the human disease data is considered a cornerstone of biomedical research, therefore, there is an enormous need for a consistent representation of human disease for robust data analysis (Colomb, 2007; LePendu et al., 2011). Creating a biomedical knowledgebase in the form of ontologies creates a rigid knowledgebase for semantic annotation of biomedical data through defined concepts and relations connecting them (Hoehndorf, Dumontier & Gkoutos, 2012). Below a group of definitions that is used as a basis for the formulation of the term ontology and its components (Fonseca, Davis & Câmara, 2003):

Definition 1. A concept in the ontology (Ψ) is a triple (τ) of the form: $\tau = [\alpha, \beta, \Omega]$, $\tau \in \Psi$, where α is the URI of the concept or a blank node, β is the predicate/property and Ω is an URI, a literal or a blank node.

Definition 2. (Semantic Relation)

A semantic relation between any two concepts in the ontology is one of the relations among the set of semantic relations $\Sigma = \{\text{Hypernym, Hyponymy, Synonymy}\}$. Semantic Relations are either taxonomic (hierarchal) or non-taxonomic relations (Ali, Fathalla, Kholief, & Hassan, 2017).

Definition 3. (Hyponym, Hypernym)

A hyponym is a word whose semantics is enclosed in another word which called its Hyperonym or Hypernym. For instance, *vaccine*, *rabies*, and *smallpox* are all Hyponyms of *viral infectious disease* (their Hypernym); that, in turn, could be a hyponym of disease by the infectious agent (Fromkin & Robert, 1998).

Definition 4. (Co-hyponyms)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

If concept C has two hyponyms A and B, then A and B are identified as co-hyponyms.

The disease ontology (DO) used in this work is a standardized biomedical ontology for describing human diseases. The objective of DO is to provide the biomedical community with a convenient, reusable and robust knowledgebase of human disease concepts, phenotype characteristics, and related medical vocabulary disease concepts. It is developed by the collaboration of different research organizations such as Northwestern University, Center for Genetic Medicine and the University of Maryland. The human disease ontology is concentrated on representing predominant and rare disease concepts captured across biomedical resources in order to provide a disease interface for human disease terminologies. Major enhancements to the DO database since 2012 has been made including: the content has had several revisions, addition of 760 terms, addition of definitions, improvement of the structure and switch to web ontology language (OWL) (Kibbe et al. 2014). The Disease Ontology database has been updated to the latest ontology as of 02-03-2017. The DO project has had a considerable influence on the development of biomedical resources, as evidenced by 291 citations of the first paper (Schriml et al., 2012) of DO. Do has imported five ontologies and contains about 33,000 classes (see Table 1).

With SPARQL, the World Wide Web Consortium standard query language for RDF, it is possible to make RDF data available through a standard interface and query it using a standard query language. The example below shows SPARQL query for finding superclasses for "Hyperuricemia" disease. SPARQL query consists basically of three parts except for the prefix declarations (lines 1–3): (a) the result specification part including solution modifiers (line 3), (b) the dataset definition part (line 4), and (c) the restriction definition part (lines 5–9) (Hartig & Heese, 2007). SPARQL query is executed against RDF triples in the RDF graph by matching these triples with criteria in the WHERE clause.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX oboInOwl:<http://www.geneontology.org/formats/oboInOwl#>
SELECT ?sc ?label ?superlabel
WHERE {
{?person rdfs:label ?label FILTER regex(str(?label),"Hyperuricemia", "i")}.
?person rdfs:subClassOf ?sc .?sc rdfs:label ?superlabel
}
```

Table 1. DO statistics

Type	Local	Imported	Inferred	Total
Number of Triples	187957	89224	2390	279453
AnnotationProperty	25	0	0	25
Class	8720	25030	0	33750
DatatypeProperty	1	0	0	1
ObjectProperty	0	1	0	1
Ontology	1	5	0	6

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The regex function is used to match two given strings. In the example above it matches the value of the variable (*?label)* with the string "*Hyperuricemia*". The purpose of the third parameter *(i)* in the regex function is that there will be case insensitive matching.

One way to access the data in the RDF model is via the Jena-embedded query engine ARQ (Seaborne, 2010) that is distributed with Jena. ARQ supports both standard SPARQL and SPARQL 1.1 for querying ontologies.

4. METHODOLOGY

The methodology of the proposed work is divided into two phases: Semantic matching and Diseases relatedness detection. In semantic matching, diseases in the user-submitted query are semantically identified in the ontology. The output of this phase is whether these diseases are found in the ontology or not. If there is a relation between them, the Uniform Resource Identifier (URI) of both diseases is retrieved. In Disease relatedness detection, the URI of each disease is forwarded to the relatedness detector to find whether they are related or not. If they are related, the algorithm returns the set of diseases that connect them in the path and the classification of both is provided as well.

4.1. Semantic Matching

The aim of this phase is to identify user-submitted diseases in the ontology graph. It is assumed that, when searching for a concept, it is also important to get synonyms of that concept. Semantic matching conducts more analysis in the matching process. The analysis is carried out to identify candidate diseases in the disease ontology (Liu, et al., 2005). A *concept matching* is obtained after disambiguation by querying WordNet lexical dictionary (Fellbaum, 1998), DO and medical datasets such as MedlinePlus (Miller, Lacroix, & Backus, 2000) and WebMD (WebMD, 2017). Each disease name in the query is first disambiguated into concepts using vector space models (Bernstein, 2005), representing concepts as vectors of features in a k-dimensional space where (k) is the number of unique relevant keywords for a given disease in the query. In other words, each disease in the query is represented by a vector of relevant keywords found through WordNet and medical datasets. Relevant keywords are *Hyponyms, Direct-Hypernyms,* and *Synonyms* of the disease. After sense disambiguation, the proposed matching algorithm returns the URIs of matched concepts.

For example, the synonyms of the disease "*Hyperuricemia*" are "*Hyperuricaemia*" and "*Uricacidemia*" therefore, diseases describing these concepts are retrieved as well. The main objective of this phase is to identify whether diseases in the query text exist in the ontology or not. The proposed algorithm iterates through ontology resources and performs semantic matching between each disease and ontology resources. When matched, sets a flag α to true and then repeat this process for the second disease. Otherwise, the disease is not found in the ontology.

4.2. Diseases Relatedness Detection

In disease relatedness detection phase, the two diseases relatedness is detected by using bidirectional spreading activation on the ontology graph which consists of several cycles including checks for termination conditions. The worthiness of bidirectional search is the speed and it requires less memory, especially

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 1. Matching algorithm

Input: Ontology O, String disease1, String disease2
Output: URI of both diseases when matched, null if not
1. α←false
2. β←false
3. ***For each*** (Resource R in O) ***do***
4. α = Semantic_Matching (R, disease1)
5. ***If*** (α≠false) ***then***
6. URI [0] = Resource.getURI(R);
7. ***End if***
8. β = semantic_matching (R, disease2)
9. ***If*** (β≠false) ***then***
10. URI [1] = Resource.getURI(R);
11. ***End if***
12. ***If*** (α≠true **AND** β≠true) ***then***
13. **Return** URI;
14. ***End if***
15. ***If*** (α≠false **OR** β≠false) ***then***
16. Print ("One of the two diseases is not found");
17. **Return** null;
18. ***End if***
End for

when dealing with large graphs such as our ontology (see Table 1) (Li & Xu, 2014). The algorithm starts with two initial nodes, i.e. the diseases concepts in the ontology graph, and follows the following steps.

1. Assume all nodes in the graph have activation values of zero. Starting nodes take an activation value of 1.0.
2. For each link $L_{i,j}$ connecting the source node η_i with target node η_j, compute $a_j = \sum_{i=1}^{n} a_i \omega_{i,j}$. where n is the number of nodes connected to η_i and ω_{ij} represents a relatedness weight, which has a value between 0.0 and 1.0.
3. If a target node receives an activation value exceeds 1.0, then set its new activation value to 1.0. Terms of the disease ontology receive the highest value of $\omega_{ij} = 1.0$ if they are one of the relevant keywords of the current activated node (i.e. hyponyms, direct-hypernyms or synonyms), because they build the taxonomy using (*is_a*) relation. A low value of $\omega_{ij} = 0.0$ is assigned to any other relations which are not always effective.
4. A firing threshold (F) determines whether to include the activated node.
5. Nodes with activation value exceeds *F* are activated.
6. Once a node has activated it will not activate again.
7. The procedure terminates when a node is reached from more than one path (i.e. *sexual disorder* in the running example shown in Figure 2).

The algorithm can identify relevant keywords by using Descriptive Logic (DL) rules in a rule base (RB). For instance, if two diseases have a relatedness type Hypernyms, then they considered the same disease. In mathematical notation, these rules defined in Table 2. These rules are formally introduced as a syntactic fragment of first-order logic. An ambiguity may arise if two homonyms are used in the search query but there is no relatedness between them. For example, "*Stewart''* in "*Stewart-Treves syndrome*",

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(a malignancy that arises within chronic lymphedema), is different from the one in "*Stewart-Bluefarb syndrome*". The latter is a type of *acro angiodermatitis* which was described independently by Stewart as well as by Bluefarb and Adams on the legs of patients with Arterio-venous malformations (Al Aboud, 2012).

By incorporating these rules, semantically matched diseases could be found using Jena reasoners. One feature of Jena is the support of different reasoners, which infer additional knowledge (Jena, Joseki & Pellet, 2004). These rules have been divided into three types: Transitive Rules, Symmetric Rules, and Co-Reflexive Rules. Jena inference system contains several inference engines (reasoners). These engines are used to check the consistency of the ontology and to infer additional RDF assertions. The primary use of this mechanism is to support the use of languages such as RDFS and OWL which allow additional facts to be inferred from instance data and class descriptions. The built-in reasoners included in Jena are (Jena, 2013):

- **Transitive Reasoner**: Supports transitive and reflexive properties traversal.
- **RDFS rule Reasoner**: Supports RDFS entailments.
- **Generic Rule Reasoner**: Supports forward chaining, backward chaining and hybrid execution strategies for user-defined rules.

The idea of a bidirectional search is to cut back the search time by looking out forward from the beginning and backward from the goal at the same time. When the two search paths meet, the algorithm constructs one path starting from the beginning node through the intersection to the goal.

5. WORKFLOW OF THE PROPOSED APPROACH

In this section, the layered architecture of the proposed method has been introduced. There are mainly three layers within the proposed conceptual architecture, namely the *Presentation Layer*, the *Middleware Layer* and the *Data Layer*. Figure 1 depicts the layered-architecture of the proposed approach.

- **Presentation Layer:** The presentation layer is responsible for providing end users the Graphical User Interface (GUI) through which users can submit queries and get the results.
- **Middleware Layer:** The primary role of the middleware layer is to provide an abstraction for ontology access. This role is implemented using Java programming language and Jena API and reasoners, which isolates clients from different API implementations and provides a unified inter-

Table 2. Rule base for the proposed method

Rule#	Mathematical Notation	Description
R_1	∀disease a, band $c \in D$: if $(aHb \wedge bHc) \rightarrow aHc$	D is the disease set in the ontology H is the "hyponym" relation.
R_2		S is the "synonym" relation.
R_3		
R_4	∀disease a, band $c \in D$: if $(aQb \wedge cQb) \rightarrow aHc$	Q is the "Hypernyms" relation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

face for querying different data models in the data layer. The second role of this layer is: 1) apply pre-processing techniques to the query, 2) formulate the semantic query, 3) query the underlying ontology and then 4) send the results back to the presentation layer.

- **Data Layer:** This layer offers access to physical data stores which is the disease ontology and medical datasets. The data stores are queried two times. One is for identifying the two diseases in the ontology, and the second is for detecting the relatedness between them.

Furthermore, SAOO prototype exemplifies how the proposed method is used. SAOO automatically processes user-submitted query in the following steps.

1. The user submits the two diseases as a string using the system interface.
2. Perform Pre-processing on input.
    a. **Intelligent Tokenization and Stop List Elimination:** This step includes tokenizing the query stream, by breaking it down into understandable segments. Then, stop words have been removed. The intelligence here is that the method does not blindly remove all stop words like traditional techniques, but using a concept tokenization technique in which the keyword is taken with the preposition after it as a concept. If that concept has a match in the ontology, then this proposition will not be removed otherwise it will be removed (Fathalla, Hassan & Kholief, 2013). Tokenization is carried out using *StringTokenizer* Java class.
    b. **Stemming:** Stemming removes word suffixes: both inflectional suffixes (*-s, -es, -ed*) and derivational suffixes (*-able, -ability*) are stemmed (Fan, Huang & An, 2006). Stemming is carried out using *Porter stemmer* (Willett, 2006) Java class.
3. Formulate semantic query using SPARQL (Jena, Joseki & Pellet, 2004) query language and executes it using the Jena-embedded query engine (ARQ) against the ontology which performs the semantic matching process described in the methodology section.
4. The diseases relatedness is detected by using bidirectional-based spreading activation on the ontology graph if the diseases found in the matching process. The output of this process is a set of diseases that builds the path between the two diseases. In addition, the classification of these diseases could be detected by performing one more cycle of the spreading process which is the disease name in one level up (if available) in the hierarchy.

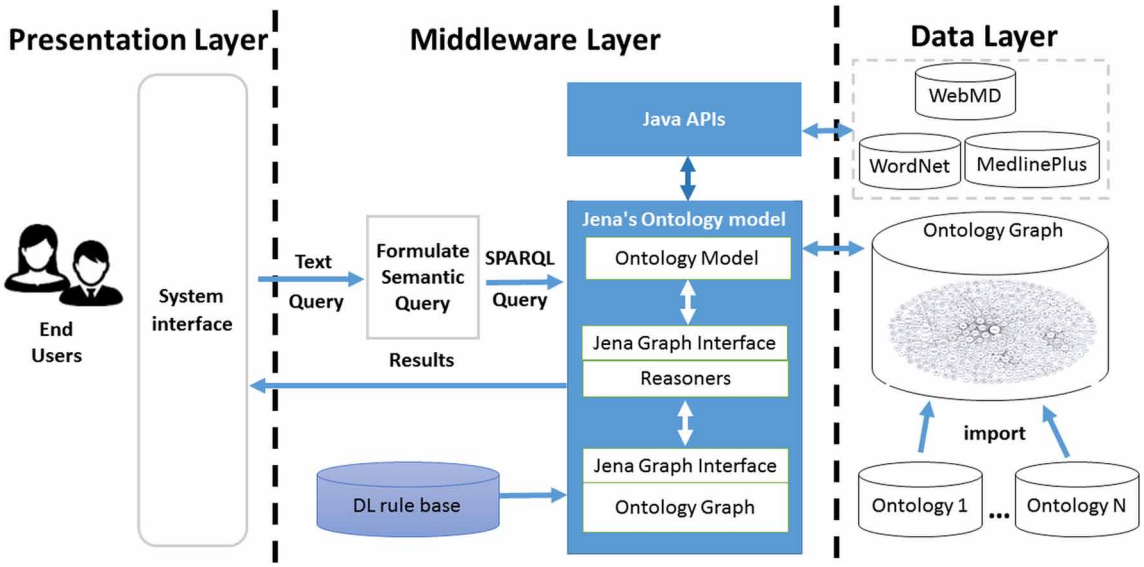
6. RUNNING EXAMPLE

As a running example in this paper, a query includes the two diseases ("*Joubert syndrome with orofaciodigital defect*" and "*Meckel syndrome*") will be used as a user-submitted query. The results of relatedness detection may consist of a set of different diseases. The relatedness between *Joubert syndrome with orofaciodigital defect* and *Meckel syndrome* diseases are detected by the intersection node *Ciliopathy*. The algorithm runs in the following steps and the whole spreading process is summarized in Table 3.

1. Initiate with two initial nodes *Joubert syndrome with orofaciodigital defect* and *Meckel syndrome* (marked in light gray).
2. In the first cycle of the activation process, one more node will be activated from each of the starting nodes, which is Joubert syndrome (from node 1) and *Ciliopathy* (from node 2).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Layered architecture of the proposed method



3. In the second cycle of the activation process, one more node will be activated from each of the latest activated nodes in each path which is *Ciliopathy* and *Genetic Disease* diseases from the first and second path respectively.
4. The algorithm detects an intersection node which is *Ciliopathy* and stops.

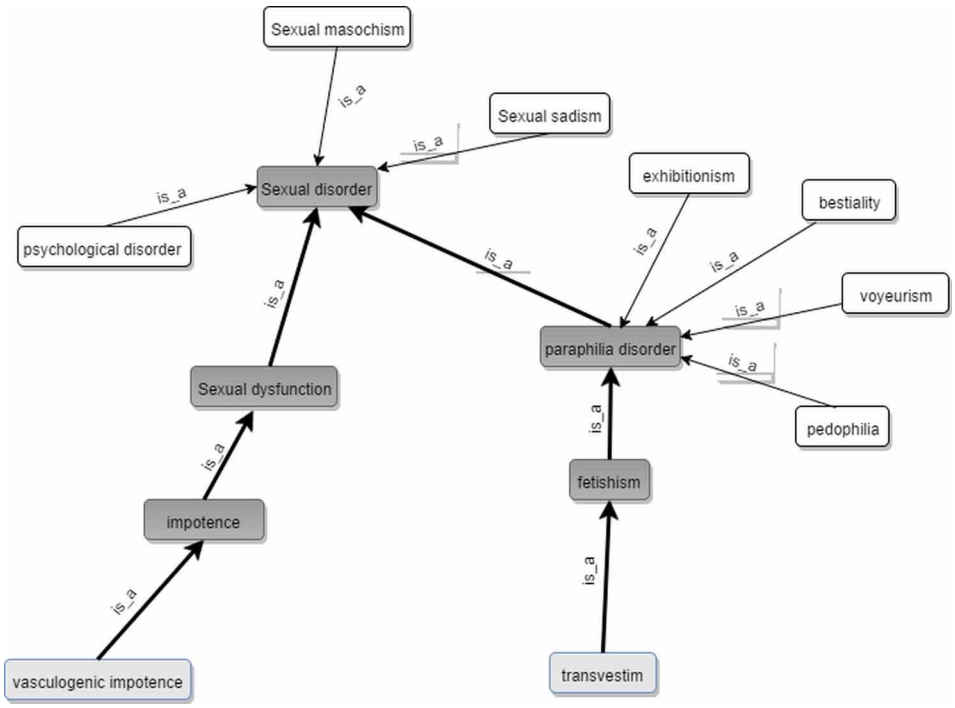
As shown in Figure 2, the algorithm originated at nodes *Vasculogenic impotence* and *Transvestism* (filled with light-grey) which has an initial activation value of 1.0. Two cycles of spreading activation have occurred. Then, they meet at the intersection node "*Sexual disorder*". Activation is spread over the graph with paths marked with bold lines and nodes filled with dark-grey. Therefore, it could be inferred that *Vasculogenic impotence* is a *sexual dysfunction,* which is *sexual disorder* and *Transvestism* is a *fetishism* which is a *paraphilia* which is also a *sexual disorder*. One more cycle can be performed to get one more up-level disease connecting both which could be considered as a classification of both diseases.

Table 3. The spreading process tracing of the running example

1st Initial Node	Joubert syndrome with orofaciodigital defect
2nd Initial Node	Meckel syndrome
Path from node 1	Joubert syndrome- Ciliopathy
Path from node 2	Ciliopathy
Intersection	Ciliopathy

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Partial view of the disease ontology shows the relatedness path (marked in bold) between "Vasculogenic impotence" and "Transvestism" diseases



7. CONCLUSION AND FUTURE WORK

In this paper, a layered architecture of the proposed method named Spreading Activation over Ontology (SAOO) has been proposed in order to detect the relatedness between two human diseases. This relatedness can be detected by running spreading activation algorithm using the bidirectional search technique on a well-formed disease ontology. One key feature of SAOO is to identify whether two human diseases are related to each other or not. Moreover, identify related diseases that may connect them in a common path. The importance of the relatedness is that the physicians can treat patients not only based on the symptoms they suffer from but treat the real cause of this disease which may be related to another disease that causes these symptoms. Therefore, the physician can treat the main cause, not just the symptoms. Consequently, SAOO helps in medical decision support. The proposed method improves the non-semantic medical systems by considering semantic domain knowledge to infer diseases relatedness. For future work, authors are planning to find the relatedness between more than two diseases using multiple-goal search algorithms and provide the relatedness graph between them not only a simple path. In addition, a relatedness set is returned to the user which is a subset of the power set of the set of input diseases. In other words, which of these diseases is related to the others? Finally, it is planned to integrate different ontologies in the medical domain using and ontology matching and integration services.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

Said Fathalla would like to acknowledge the Ministry of Higher Education (MoHE) of Egypt for providing a scholarship to conduct this study as well as the University of Bonn, Germany for offering all facilities and tools required to do this work.

REFERENCES

Al Aboud, A., & Al Aboud, K. (2012). Similar names and terms in dermatology; an appraisal. *Our Dermatol Online*, *3*(4), 367–368. doi:10.7241/ourd.20124.86

Ali, A. B. M. (Ed.). (2009). *Dynamic and Advanced Data Mining for Progressing Technological Development: Innovations and Systemic Approaches*. IGI Global.

Ali, M., Fathalla, S., Kholief, M., & Hassan, Y. F. (2017, September). The problem learning Non-Taxonomic Relationships of Ontologies from unstructured data sources. In *Proceedings of the 2017 23rd International Conference on Automation and Computing (ICAC)*. IEEE. 10.23919/IConAC.2017.8082083

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. doi:10.1016/S0022-5371(83)90201-3

Antoniou, G., & Harmelen, F. (2004). *A semantic web primer*. MIT press.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*(5), 28–37. doi:10.1038cientificamerican0501-34 PMID:11341160

Bernstein, A., Kaufmann, E., Burki, C., & Klein, M. (2005). How Similar Is It? Towards Personalized Similarity Measures in Ontologies. In *Proceedings of the 7th International Conference Wirtschafts informatik (WI-2005)*, Bamberg, Germany. 10.1007/3-7908-1624-8_71

Burns, E. C., Tanner, J. M., Preece, M. A., & Cameron, N. (1981). Final height and pubertal development in 55 children with idiopathic growth hormone deficiency, treated for between 2 and 15 years with human growth hormone. *European Journal of Pediatrics*, *137*(2), 155–164. PMID:7308226

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. doi:10.1037/0033-295X.82.6.407

Colomb, R. M. (2007). *Ontology and the Semantic Web*. In Frontiers in Artificial Intelligence and Applications (Vol. 156).

Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., ... Jassal, B. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Research*, *42*(D1), D472–D477. doi:10.1093/nar/gkt1102 PMID:24243840

De Maio, C., Loia, V., Fenza, G., Gallo, M., Linciano, R., & Morrone, A. (2011, June). Fuzzy knowledge approach to automatic disease diagnosis. In *Proceedings of the 2011 IEEE International Conference on Fuzzy Systems (FUZZ)* (pp. 2088-2095). IEEE. 10.1109/FUZZY.2011.6007498

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Dhamankar, R., Lee, Y., Doan, A., Halevy, A., & Domingos, P. (2004, June). iMAP: discovering complex semantic matches between database schemas. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data* (pp. 383-394). ACM. 10.1145/1007568.1007612

Do, H. H., & Rahm, E. (2002, August). COMA: a system for flexible combination of schema matching approaches. In *Proceedings of the 28th international conference on Very Large Data Bases* (pp. 610-621). 10.1016/B978-155860869-6/50060-3

Euzenat, J., Meilicke, C., Stuckenschmidt, H., Shvaiko, P., & Trojahn, C. (2011). Ontology alignment evaluation initiative: Six years of experience. *Journal on Data Semantics*, *15*, 158–192.

Fan, Y., Huang, X., & An, A. (2006). York University at TREC 2006: Enterprise Email Discussion Search.

Fathalla, S., & Kannot, Y. (2017, September). A Bidirectional-Based Spreading Activation Method for Human Diseases Relatedness Detection Using Disease Ontology. In *Proceedings of the Conference on Computational Collective Intelligence Technologies and Applications* (pp. 14-23). Springer, Cham. 10.1007/978-3-319-67074-4_2

Fathalla, S., & Kannot, Y. M. K. (2017). Bidirectional Spreading Activation Method for Finding Human Diseases Relatedness Using Well-Formed Disease Ontology. *International Journal of Computers in Clinical Practice*, *2*(1), 45–58. doi:10.4018/IJCCP.2017010104

Fathalla, S., Vahdati, S., Auer, S., & Lange, C. (2017, September). Towards a Knowledge Graph Representing Research Findings by Semantifying Survey Articles. In *Proceedings of the International Conference on Theory and Practice of Digital Libraries* (pp. 315-327). Springer, Cham. 10.1007/978-3-319-67008-9_25

Fathalla, S. M., Hassan, Y. F., & El-Sayed, M. (2012, October). A hybrid method for user query reformation and classification. In *Proceedings of the 2012 22nd International Conference on Computer Theory and Applications (ICCTA)* (pp. 132-138). IEEE. 10.1109/ICCTA.2012.6523559

Fathalla, S. M., Hassan, Y. F., & Kholief, M. (2013). Ontology-Based Approach for Automated Issue Classification in an Issue Tracking System. In *Proceedings of the 23rd International Conference on Computer Theory and Applications (ICCTA)*. IEEE.

Fellbaum, C. (1998). WordNet an Electronic Lexical Database. *Computational Linguistics*, *25*(2), 292–296.

Fonseca, F., Davis, C., & Câmara, G. (2003). Bridging ontologies and conceptual schemas in geographic information integration. *GeoInformatica*, *7*(4), 355–378. doi:10.1023/A:1025573406389

Fromkin, V., & Robert, R. (1998). *Introduction to Language* (6th ed.). Fort Worth: Harcourt Brace College Publishers.

Giunchiglia, F., Shvaiko, P., & Yatskevich, M. (2004, May). S-Match: an algorithm and an implementation of semantic matching. In *European semantic web symposium* (pp. 61-75). Springer Berlin Heidelberg.

Giunchiglia, F., Yatskevich, M., & Shvaiko, P. (2007). Semantic matching: Algorithms and implementation. *Journal on Data Semantics*, *9*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gonzalo, J., Li, H., Moschitti, A., & Xu, J. (2014, July). SIGIR 2014 workshop on semantic matching in information retrieval. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval* (pp. 1296-1296). ACM.

Guo, J., Fan, Y., Ai, Q., & Croft, W. B. (2016, October). Semantic Matching by Non-Linear Word Transportation for Information Retrieval. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management* (pp. 701-710). ACM. 10.1145/2983323.2983768

Hartig, O., & Heese, R. (2007, June). The SPARQL query graph model for query optimization. In *Proceedings of the European Semantic Web Conference* (pp. 564-578). Springer Berlin Heidelberg. 10.1007/978-3-540-72667-8_40

Hasan, S., & Curry, E. (2014). Approximate semantic matching of events for the internet of things. *ACM Transactions on Internet Technology*, *14*(1), 2. doi:10.1145/2633684

He, Y., Cowell, L., Diehl, A. D., Mobley, H. L., Peters, B., Ruttenberg, A., ... Xiang, Z. (2009, July). VO: vaccine ontology. In *Proceedings of the 1st International Conference on Biomedical Ontology (ICBO 2009)*, Buffalo, NY.

Hoehndorf, R., Dumontier, M., & Gkoutos, G. V. (2012). Evaluation of research in biomedical ontologies. *Briefings in Bioinformatics*. PMID:22962340

Jena, A. (2013). *Reasoners and rule engines: Jena inference support*. The Apache Software Foundation.

Joseki, J., & Pellet, F. (2014). Semantic web frameworks. Oaks, R., University of Sterling. Retrieved from http://www.cs.stir.ac.uk/courses/31Z7/posters/2014/rao.pdf

Kacholia, V., Pandit, S., Chakrabarti, S., Sudarshan, S., Desai, R., & Karambelkar, H. (2005, August). Bidirectional expansion for keyword search on graph databases. In *Proceedings of the 31st international conference on Very large data bases* (pp. 505-516). VLDB Endowment.

Kanavos, A., Theodoridis, E., & Tsakalidis, A. (2014, September). A pubmed meta search engine based on biomedical entity mining. In *Proceedings of the 2014 25th International Workshop on Database and Expert Systems Applications (DEXA)* (pp. 82-86). IEEE. 10.1109/DEXA.2014.32

Kibbe, W. A., Arze, C., Felix, V., Mitraka, E., Bolton, E., Fu, G., ... Parkinson, H. (2014). Disease Ontology 2015 update: An expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Research*. PMID:25348409

Kohler, S., Doelken, S. C., Mungall, C. J., Bauer, S., Firth, H. V., Bailleul-Forestier, I., ... Robinson, P. N. (2014). The Human Phenotype Ontology project: Linking molecular biology and disease through phenotype data. *Nucleic Acids Research*, *4*(D1), 966–974. doi:10.1093/nar/gkt1026 PMID:24217912

LePendu, P., Musen, M. A., & Shah, N. H. (2011). Enabling enrichment analysis with the human disease ontology. *Journal of Biomedical Informatics*, *44*, S31–S38. doi:10.1016/j.jbi.2011.04.007 PMID:21550421

Li, H., & Xu, J. (2014). Semantic matching in search. *Foundations and Trends in Information Retrieval*, *7*(5), 343-469.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, W., Weichselbraun, A., Scharl, A., & Chang, E. (2005). Semi-automatic ontology extension using spreading activation. *Journal of Universal Knowledge Management*, *1*(1), 50–58.

Miller, N., Lacroix, E. M., & Backus, J. E. (2000). MEDLINEplus: Building and maintaining the National Library of Medicine's consumer health Web service. *Bulletin of the Medical Library Association*, *88*(1), 11. PMID:10658959

Ngo, V. M., Cao, T. H., & Le, T. M. (2010, November). Combining named entities with wordnet and using query-oriented spreading activation for semantic text search. In *Proceedings of the 2010 IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*. IEEE. 10.1109/RIVF.2010.5633401

Oberle, D. (2014). How ontologies benefit enterprise applications. *Semantic Web*, *5*(6), 473–491.

Qin, Y., Yao, L., & Sheng, Q. Z. (2016, September). Approximate Semantic Matching Over Linked Data Streams. In *Proceedings of the International Conference on Database and Expert Systems Applications* (pp. 37-51). Springer International Publishing. 10.1007/978-3-319-44406-2_5

Schriml, L. M., Arze, C., Nadendla, S., Chang, Y. W., Mazaitis, M., Felix, V., ... Kibbe, W. A. (2012). Disease Ontology: A backbone for disease semantic integration. *Nucleic Acids Research*, *40*(D1), D940–D946. doi:10.1093/nar/gkr972 PMID:22080554

Schumacher, K., Sintek, M., & Sauermann, L. (2008, June). Combining fact and document retrieval with spreading activation for semantic desktop search. In *Proceedings of the European Semantic Web Conference* (pp. 569-583). Springer Berlin Heidelberg. 10.1007/978-3-540-68234-9_42

Seaborne, A. (2010). ARQ-A SPARQL Processor for Jena. Retrieved from http://jena.sourceforge.net/ARQ

Shadbolt, N., Berners-Lee, T., & Hall, W. (2006). The semantic web revisited. *IEEE Intelligent Systems*, *21*(3), 96–101. doi:10.1109/MIS.2006.62

Shvaiko, P., & Euzenat, J. (2013). Ontology matching: State of the art and future challenges. *IEEE Transactions on Knowledge and Data Engineering*, *25*(1), 158–176. doi:10.1109/TKDE.2011.253

WebMD. (n.d.). Retrieved January 2017 from http://www.webmd.com/

Willett, P. (2006). The Porter stemming algorithm: Then and now. *Program*, *40*(3), 219–223. doi:10.1108/00330330610681295

Wu, Z., Zhu, H., Li, G., Cui, Z., Huang, H., Li, J., ... Xu, G. (2017). An efficient Wikipedia semantic matching approach to text document classification. *Information Sciences*, *393*, 15–28. doi:10.1016/j.ins.2017.02.009

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] Pubmed: http://www.ncbi.nlm.nih.gov/pubmed
[2] MeSH: https://www.ncbi.nlm.nih.gov/mesh
[3] RxNorm: https://www.nlm.nih.gov/research/umls/rxnorm/
[4] Disease Ontology: http://www.disease-ontology.org
[5] Vaccine Ontology (VO): http://www.violinet.org/vaccineontology
[6] Infectious Disease Ontology: http://infectiousdiseaseontology.org/page/Main_Page
[7] Ontology for Biomedical Investigations: http://obi-ontology.org/page/Main_Page
[8] Ontology for General Medical Science: https://bioportal.bioontology.org/ontologies/OGMS

This research was previously published in the International Journal on Semantic Web and Information Systems (IJSWIS), 14(3); edited by Miltiadis D. Lytras; pages 120-133; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 27
An Artificial Neural Network Classification of Prescription Nonadherence

Steven Walczak
University of South Florida, USA

Senanu R. Okuboyejo
Covenant University, Nigeria

ABSTRACT

This study investigates the use of artificial neural networks (ANNs) to classify reasons for medication nonadherence. A survey method is used to collect individual reasons for nonadherence to treatment plans. Seven reasons for nonadherence are identified from the survey. ANNs using backpropagation learning are trained and validated to produce a nonadherence classification model. Most patients identified multiple reasons for nonadherence. The ANN models were able to accurately predict almost 63 percent of the reasons identified for each patient. After removal of two highly common nonadherence reasons, new ANN models are able to identify 73 percent of the remaining nonadherence reasons. ANN models of nonadherence are validated as a reliable medical informatics tool for assisting healthcare providers in identifying the most likely reasons for treatment nonadherence. Physicians may use the identified nonadherence reasons to help overcome the causes of nonadherence for each patient.

INTRODUCTION

Prescription (pharmaceutical) treatment plans require a specific pharmaceutical drug to be taken in specific amounts at specific intervals for a specific period of time. Nonadherence is the violation of any of these specified treatment requirements: not taking the correct dosage (too little or too much), missing or delaying scheduled administrations of the pharmaceutical, or not completing the treatment (Hugtenburg et al., 2013). Nonadherence to pharmaceutical treatment plans is a worldwide dilemma and prior research results examining the rate of nonadherence in 20 countries are shown in Table 1. As

DOI: 10.4018/978-1-7998-1204-3.ch027

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

may be seen, documented nonadherence rates vary in different cultures, but range from 10% to 88%. Prescription nonadherence is a persistent problem in healthcare today (Lehane & McCarthy, 2007) and nonadherence through underuse of a medication is rising significantly in prevalence (Kirking et al., 2006).

Nonadherence, especially to antibiotics and antimalarials, causes increased risks to the population by enabling evolution of resistant strains of malaria and other diseases (Andrajati et al., 2016; Awad & Eltayeb, 2007; Okuboyejo, 2014; Pechère, 2001) and harms population health by promoting further spread of diseases due to resulting ineffective treatment (Andrajati et al., 2016; Awad & Eltayeb, 2007; Center for Disease Control, 2013; Gibson et al., 2011; Okuboyejo & Eyesan, 2014). The Center for Disease Control states that pharmaceutical treatment nonadherence leads to significant economic and well-being impacts with direct costs estimated at $289 billion annually and 125,000 deaths annually (Center for Disease Control, 2013).

Reasons for patient nonadherence with the pharmaceutical portion of their treatment plans vary. Self-medication is a rising worldwide problem (Agarwal, Yewale, & Dharmapalan, 2015) especially in regions of the world where antibiotics and other medications are available without prescription, and has been identified as a contributing factor to improper treatment including nonadherence to recommended treatment plans (Awad & Eltayeb, 2007; Grigoryan et al., 2006; Zhu et al., 2016). Other reasons for pharmaceutical nonadherence are: cost of medication (Center for Disease Control, 2013; Gibson et al., 2011; Hirth et al., 2012; Kirking et al., 2006), negative attitude toward drug or don't like taking drugs (Kirking et al., 2006; Okuboyejo, 2014; Urquhart, 2005), lack of time/employment/travel (Au et L., 2014; Okuboyejo, 2014), don't trust or other issues with physician (Okuboyejo, 2014, Zhu et al., 2016), lack of (Kirking et al., 2006; Urquhart, 2005), forgetfulness (Au et al., 2014), religious or cultural reasons (Urquhart, 2005), fear of side effects (Gottlieb, 2000; Kirking et al., 2006), and other personal reasons (Urquhart, 2005). It is important to note that each patient may have multiple reasons for being nonadherent (Okuboyejo, 2014; Urquhart, 2005).

Various methods have been proposed to lessen nonadherence, including: better physician-patient communication (Agarwal, Yewale, & Dharmapalan, 2015; Buckalew & Sallis, 1986; Okuboyejo, 2014), other patient education (Oyelami, Okuboyejo, & Ebiye, 2013), and sending automatic reminders (Eyesan & Okuboyejo, 2013; Okuboyejo, Omoregbe, & Mbarika, 2012). Each of these proposed solutions is generic for a whole population of patients and typically addresses only a single reason for nonadherence. Identifying an individual patient's specific reasons for nonadherence would enable tailored individualized responses to address the reasons for nonadherence relevant to each patient. An artificial neural network (ANN) informatics tool to predict the nonadherence reasons for individual patients is developed based on Social Learning Theory. The ANN is tested for predicting individual nonadherence reasons on a population of patients in Nigeria.

BACKGROUND

Brief Background on Artificial Neural Networks

ANNs are nonparametric machine learning systems inspired by the human brain and composed of processing elements arranged in layers and highly interconnected (see Figure 1) (Caudill, 1987). Each connection is initially assigned a randomized value or weight. Input layer nodes receive raw numeric values representing the independent variables. Additional processing nodes are arranged into some number of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

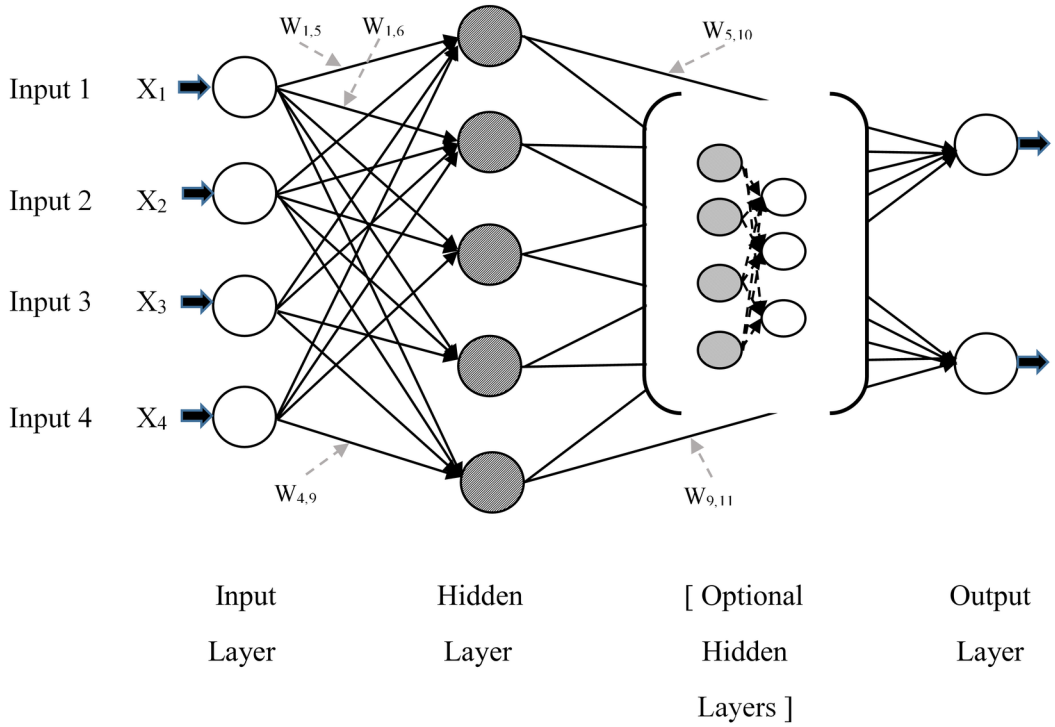
Table 1. Documented nonadherence rates from around the world

Country	Nonadherence rates	Reference
Australia/New Zealand	16.4%	(Hirth et al., 2012)
Belgium	12% 18%	(Hirth et al., 2012) (Pechère, 2001)
Canada	15%	(Hirth et al., 2012)
China	57-64%	(Zhu et al., 2016)
Columbia	41%	(Pechère, 2001)
France	12.3% 18%	(Hirth et al., 2012) (Pechère, 2001)
Germany	12.9%	(Hirth et al., 2012)
India	17%	(Agarwal, Yewale, & Dharmapalan, 2015)
Indonesia	43%	(Andrajati et al., 2016)
Italy	13.1% 35%	(Hirth et al., 2012) (Pechère, 2001)
Japan	10.9%	(Hirth et al., 2012)
Malaysia	29%	(Agarwal, Yewale, & Dharmapalan, 2015)
Morocco	38%	(Pechère, 2001)
Spain	12.7% 45%	(Hirth et al., 2012) (Pechère, 2001)
Sweden	11.8%	(Hirth et al., 2012)
Thailand	47%	(Pechère, 2001)
Turkey	30%	(Pechère, 2001)
United Kingdom	10% 13.2%	(Pechère, 2001) (Hirth et al., 2012)
United States	7% - 45% 18% 20 – 80% 38% 45% – 88% 50% 56.2%	(Kirking et al., 2006) (Hirth et al., 2012) (Gottlieb, 2000) (Au et al., 2014) (Gamble, 2009) (Buckalew & Sallis, 1986; Wapner, 2008) (Gibson et al., 2011)
Worldwide	0% - 95.4%	(DiMatteo, 2004)

hidden processing layers and finally an output layer, which represents the dependent variable(s). All hidden and output nodes receive the aggregation of the product of the incoming nodes and the associated connection weights, thus if summation is used as the aggregation function then a node's input value is equal to $\Sigma\, w_{ij}x_i$, where w_{ij} is the weight on the connection from node *i* to node *j*, coming from node x_i. Each processing element performs its own computation on this input value, known as a transfer function, before transmitting it along its outbound connections. Transfer and aggregation functions may be any function defined by the user, such as sum or average or maximum for the aggregation function and sigmoid or a hyperbolic tangent function for the transfer function.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Artificial neural network architecture (only a few connection weights shown for clarity)



The backpropagation learning algorithm, learns by propagating the error term of the output produced by a given set of input against the known values backwards through the network to adjust the connection weights. Backpropagation is known as a supervised learning system since it requires training records with known outcomes in order to calculate the correct output error.

One advantage of ANNs over other modeling paradigms is that since they are nonparametric they may work with any type of numeric input values and those values have no constraints as in other statistical models, such as requiring normal distribution or heteroskedastic error. Translating data into numeric representations for use in an ANN is easy to perform, such as representing a categorical variable with a corresponding number of input variables (one for each category) that are set to true or false. Second, since ANNs are learning systems, they are able to find optimal solution surfaces to model the given training data. This includes complex nonlinear solution surfaces. Additionally, once trained, ANNs are highly tolerant of noise in future input data. Disadvantages are that ANNs are sensitive to the independent variables selected, which must not be highly correlated, and the number of nodes in each hidden layer, which can cause poor out of sample performance due to memorization (Smith, 1993; Walczak & Cerpa, 1999). These disadvantages may be overcome by first testing desired independent variables with a correlation test, such as Pearson's Correlation Matrix, and by developing and testing multiple hidden node and layer architectures to find the optimal quantity of layers and processing nodes.

A more detailed explanation of ANNs is beyond the scope of this paper. Many tutorials exist and readers are encouraged to pursue these for additional explanation (see (Jain, Mao, & Mohiuddin, 1996; Rodvold, McLeod, & Brandt, 2001; Walczak & Cerpa, 1999; Zhang, 2007) for examples, guidelines, and further information on ANN development).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Prescription Nonadherence Research

As shown in Table 1, research on prescription nonadherence has been conducted that examines populations from around the world for at least the last sixty years (DiMatteo, 2004). Most of the reported research consists of empirical studies on specific populations for a single or small set of medical reasons (DiMatteo, 2004). Different theoretical models commonly used in prescription nonadherence research include: the Health Belief Model, Social Learning Theory, Theory of Planned Behavior, and the Transtheoretical Model (Bowen, Helmes, & Lease, 2001; DiMatteo & Martin, 2002).

Prescription nonadherence research uses multiple direct and indirect methods in evaluating nonadherence (Farmer, 1999). Indirect methods include patient self-reporting and pill counting. The research reported in this article uses self-reporting. Direct methods involve urine or blood tests to detect the presence of the desired amount of the prescribed pharmaceutical.

Despite the plethora of research on prescription nonadherence, results have not been either cost effective or clinically effective in reducing nonadherence (Lehane & McCarthy, 2007). Several systematic studies have been performed over this large body of existing literature (DiMatteo, 2004; Hayes et al., 2002; McDonald, Garg, & Haynes, 2002) and commonly classify reasons for nonadherence as either intentional factors including physician-patient communication or unintentional factors such as forgetfulness (Lehane & McCarthy, 2007).

METHOD

Population and Data Collection

A survey based on Akers (1977) Social Learning Theory is designed using prior surveys examining pharmaceutical treatment nonadherence in Africa (Ekwunife, Udeogaranya, & Adibe, 2010; Erah & Arute, 2008; Ikechuwku, Obinna, & Ogochukwu, 2010; Uzochukwu et al., 2009). The survey contained five demographic questions (age, sex, marital status, education level, and payer) and 25 questions covering intentional and unintentional reasons for being nonadherent.

After obtaining informed consent and guaranteeing anonymity, the survey is given to 600 outpatients from university and government hospitals and clinics in southwest Nigeria who were taking any type of pharmaceutical as part of their prescribed treatment. A total of 474 responses are returned for a response rate of almost 79 percent.

Data Preparation and Independent Variable Definition

The anonymized raw data is then transcribed into a spreadsheet format for analysis and preparation to further format the data for the ANN. Ten of the responses are discarded because the respondents indicated they were not currently on a pharmaceutical treatment plan. Since respondents were instructed that they didn't have to answer any question which made them feel uncomfortable, numerous responses were missing values. Whether the missing values are due to discomfort or other reasons including laziness is unknown. Further preparation of the response data for use by the ANNs eliminates responses with too many missing values (Cherkassky & Lari-Najafi, 1992), leaving a total of 237 usable responses. The usable responses are divided into a training set of 137 responses and a validation set or hold-out sample

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of 100 responses. Three-fold cross validation is used to enable each response to serve as both a training and validation sample, but for different training instances of each ANN. The validation samples are never seen during the ANN training and can thus simulate real-world usage of the ANN for predicting individual patient reasons for nonadherence. Each response is only counted as a validation item a single time.

The survey specifies 23 specific reasons for nonadherence behavior and respondents are able to select as many as apply to them. These 23 individual reasons are further classified into seven distinct classifications of reasons for nonadherence which are: a lack of understanding or education, issues with the doctor (e.g., distrust or dislike), cost of the medication, religious objections, concern over side effects, lack of time or other time issues, and other personal reasons. Lack of understanding or education, time issues, and other personal reasons fall under unintentional reasons, while the remaining four categories of nonadherence reasons are intentional. These seven categories serve as output variables for the ANN.

ANN Design and Development

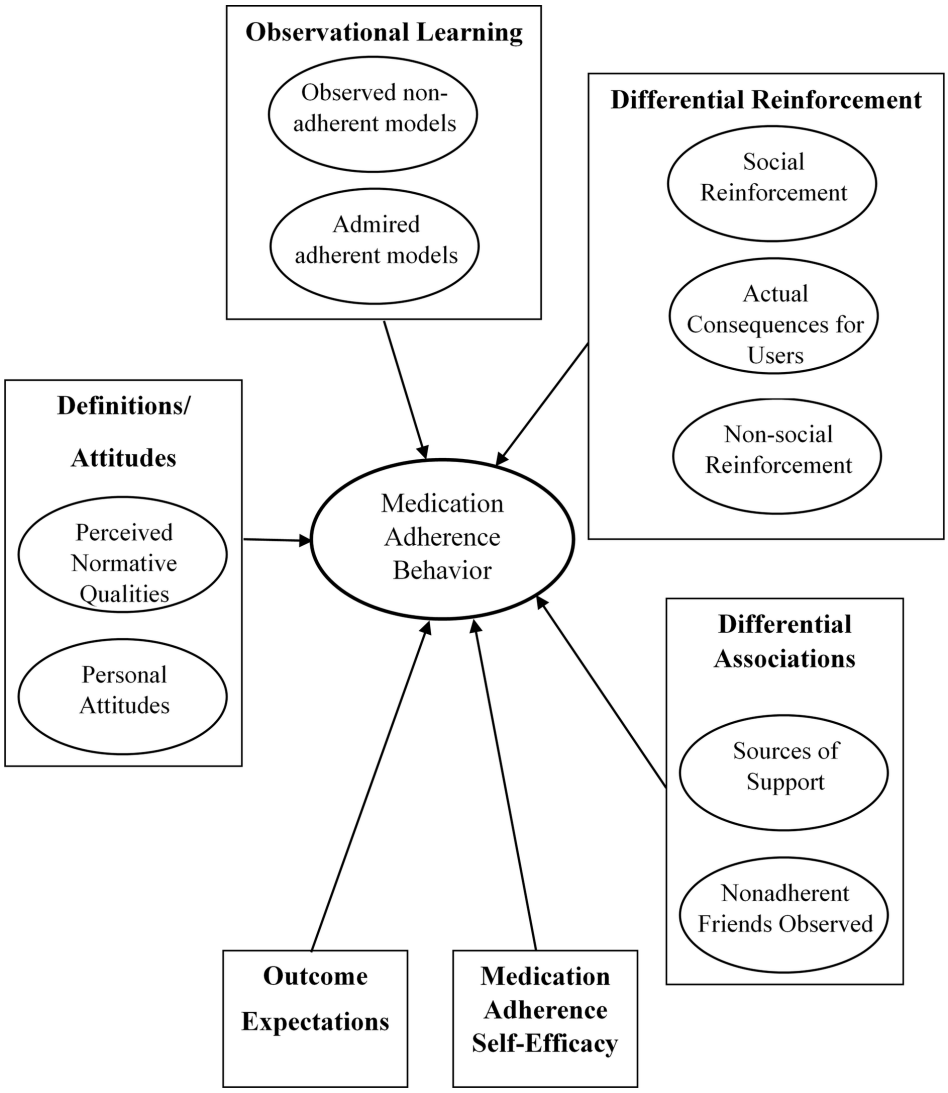
The demographic variables age, marital status, sex, education, and primary source of payment (self, family, work, insurance, charity, other) and 25 additional variables representing observational learning, differential association, differential reinforcement, medication adherence self-efficacy, outcome expectations, and personal attitudes towards adherence that are proposed influences on adherent behavior, as displayed in Figure 2. This creates a 34 variable input vector, since some of the demographic variables are treated as categorical variables.

Various training methods including radial basis function (RBF), traditional backpropagation or multi-layer perceptron, and backpropagation with two-layer forward connections are evaluated for classifying nonadherence reasons. The backpropagation with forward connections outperformed the other two learning methods and is the only one discussed for the remainder of this article. Backpropagation is the most common learning and training method used in ANNs and because of its robustness and ease of implementation and enabling easier comparison with future ANN research in nonadherence (Given, Given, & Simoni, 1978). Both single and two hidden layer architectures are investigated. The single layer architectures performed at least as well as the two hidden layer architectures and for simplicity (reducing the total number of connections in the ANN), the single hidden layer models were selected as the best performing architecture. The best performing ANN architecture had 34 input values, 42 hidden layer nodes in a single hidden layer, and seven output nodes. Using just a single hidden layer is informative as it indicates the solution surface is a hyperplane with a single concavity or convexity with no additional nonlinearities in the surface (Walczak & Cerpa, 1999).

After finding a single hidden layer ANN architecture that produced the best classification performance for likely nonadherence reasons for individual patients, further examination of the data revealed that less than three percent of individuals indicating cost was a reason had someone other than themselves or their family paying for the medication. Almost a third of those that self-paid had concerns about cost. Additionally, 85 percent of the respondents indicated that lack of education was a reason for nonadherence, but this answer may have been biased since a university was conducting the survey and this type of answer would seem desirable for a university research outcome. Therefore, if we assume that cost is usually an issue and that good communication between the physician and patient to educate them about the benefits of their medication is needed, then a new ANN model may be developed examining the remaining five nonadherence reasons for individual patients. Additionally, instead of using the 25 elements for the six factors influencing adherence shown in Figure 2, these are represented by the individual

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Factors influencing medication adherence



factors (two were used for attitudes and definitions). The best performing architecture for this new ANN is 13 input nodes, 11 hidden nodes in a single layer, and five output nodes as shown in Figure 3. This new ANN reduces the total number of connections in the best performing ANN from 1722 to 198 and thus reduces the chance for overfitting the data and consequent improvement to generalization of the results.

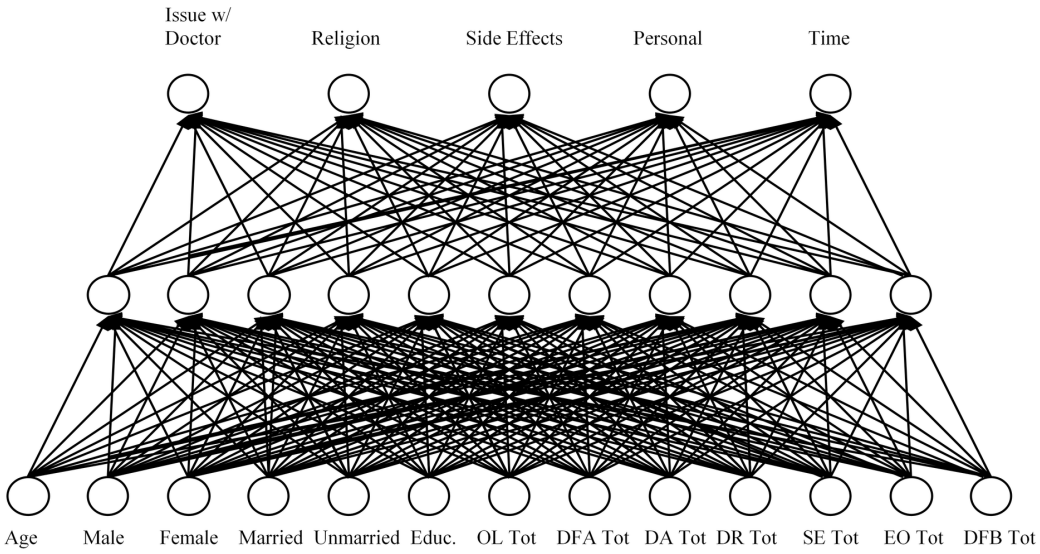
RESULTS

Population Demographic Results

Respondent age is captured in 10 year increments with a minimum of 18 (18-27) and a maximum of 67 (58-67). The median and mode for age are both the 28-37 year old group. Fifty percent of the respondents

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Backpropagation ANN architecture for predicting reasons for prescription nonadherence (not all connections shown)



were married and 44.8 percent were single or divorced or widowed, with the remaining having a missing value. Males represented 46 percent and females 47 percent of the population with the remaining having a missing value. Three quarters were employed or self-employed, with 18.2 percent reporting unemployment and 6.8 percent unknown.

ANN Prediction Results

The resulting ANN system is called INReP (Individual Nonadherence Reason Predictor). Results for the INReP ANN predicting all 7 categories of nonadherence are shown in Table 2. The INReP ANN did best predicting when religious concerns were or were not a nonadherence reason. For every nonadherence reason, INReP was correct over 54 percent of the time. The results may be further analyzed with results to specificity and sensitivity. A total of 550 true positive predictions (where positive indicates the respondent indicated this was a reason for their nonadherence), 493 true negatives, 132 false positives, and 128 false negatives resulted, producing a sensitivity to nonadherence reasons of 64.5 percent and a specificity to lack of nonadherence reasons of 61.2 percent.

The INReP ANN that predicted the 5 class subset of nonadherence reasons is shown in Table 3. A total of 479 true positive predictions, 384 true negatives, 197 false positives, and 125 false negatives resulted, producing a sensitivity to nonadherence reasons of 79.3 percent and a specificity to lack of nonadherence reasons of 66.1 percent.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. INRePP ANN results for predicting up to seven possible reasons for nonadherence

Nonadherence Category	INRePP ANN correct predictions* ($N = 237$)
Education/Understanding	175 (73.8%)
Issues with Doctor	135 (57.0%)
Cost	138 (58.2%)
Religion	183 (77.2%)
Concern over Side Effects	152 (64.1%)
Time	130 (54.9%)
Other Personal	130 (54.9%)
Total	1043 (62.9%)

*A prediction is considered correct if it was either a true positive or a true negative

Table 3. INRePP ANN results for predicting only 5 reasons for nonadherence (minus education and cost)

Nonadherence Category	INRePP ANN correct predictions* (N = 237)
Issues with Doctor	166 (70.0%)
Religion	197 (83.1%)
Concern over Side Effects	175 (73.8%)
Time	147 (62.0%)
Other Personal	178 (75.1%)
Total	863 (72.8%)

*A prediction is considered correct if it was either a true positive or a true negative

DISCUSSION

As shown above, the two INRePP ANN models both have reasonably high sensitivity for predicting the reasons a specific patient will become nonadherent and specificity for eliminating individual reasons of nonadherence. Naturally, assuming better education is always required and will improve medication adherence (Erah & Arute, 2008) and that cost is commonly an issue for self-payers, produces better prediction results, specifically an almost 23 percent improvement in sensitivity and an 8 percent improvement in specificity.

In addition to promoting increased patient communication with their physicians, other means may provide alternatives for patients to become better informed about their drugs and the need to complete treatment. The Internet is becoming much more widely used in Nigeria and provides a valuable resource for patient education (Oyelami, Okuboyejo, & Ebiye, 2013). Education can help address multiple nonadherence reasons beyond education or understanding, including concern over side effects, overcoming observational learning and social reinforcement effects, and personal reasons associated with dislike of medicine or concerns over drug habits.

Methods for overcoming lack of time or the personal reason of forgetfulness have been recommended in the literature. These include placing automated telephone calls to patients (Eyesan & Okuboyejo, 2013) and sending text messages (Okuboyejo, Omoregbe, & Mbarika, 2012) to remind patients of current drug

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

delivery times. Patients may also use or be provided alarm mechanisms (e.g., watch or cell phone) and carrying cases to assist them with remembering prescription administration time and enabling them to have their prescriptions readily available, though such mechanisms will increase the overall cost of care.

A need for future research to examine embedding ANN decision support into electronic heath records (EHRs) is still needed to make the results of the ANNs more readily available to physicians when getting ready to prescribe a pharmaceutical-based treatment plan. EHRs have been shown to improve patient safety within hospitals while enabling the improvement in the quality of care delivered in hospital setting (Bourgeois & Yaylacicegi, 2010). Patient safety and outcome concerns however, do not end at the hospital door. Patients being treated pharmacologically, are commonly monitored for adverse drug effects (Gupta et al., 2007), but this needs to be extended to assisting patients to follow and complete their prescription treatment courses to the fullest in order to obtain desired health outcomes.

Decision support regarding dangerous drug interactions already exist in EHRs (Helmons et al., 2015) and including the INReP ANN prediction of nonadherence reasons along with other drug related decision support should be a simple extension to existing EHRs (Kuperman et al., 2007). Research has shown that when EHRs are available 65 percent of physicians use them to examine drug interactions and to get recommendations when prescribing pharmaceuticals (DeMello & Deshpande, 2012). With the INReP predictions embedded in the EHR utilized by a healthcare practice or hospital, physicians can optimize patient outcomes by including education, reminders, or other initiatives to overcome individual patients' reasons for nonadherence, thus improving outcomes not only for the individual patient, but for the population health as well.

Another area for further research, but outside of the scope of information systems research is that of response bias. Recall that while the survey had a reasonable response rate of 79 percent, only half of these responses could be used in the ANN research due to missing values. Missing data values increases the uncertainty related to the data and nay cause to decrease the subsequent reliability of various modeling paradigms, including ANNs. Further research should investigate the reasons why respondent elect to not answer all of the questions, especially values critical to the research, in medically oriented domains to determine if further explanation or guarantees of anonymity could alleviate this problem.

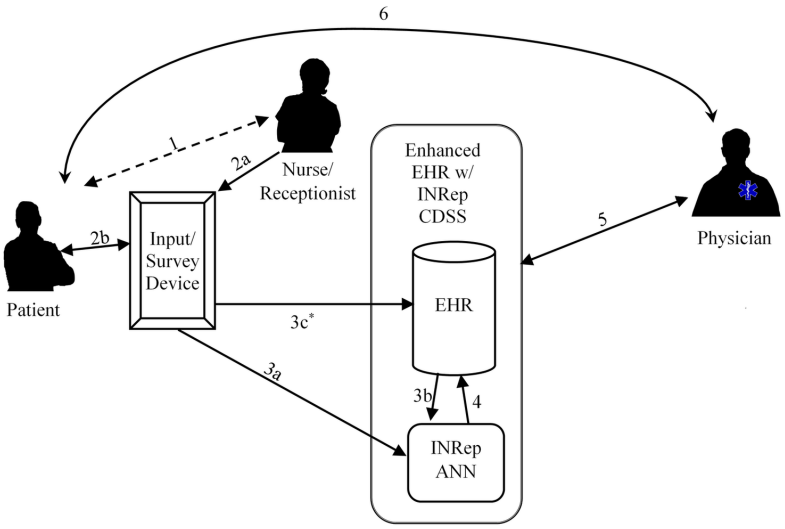
A short questionnaire provided to patients during check-in or triage would produce the values needed by the ANN clinical decision support informatics tool for predicting nonadherence reasons for individual patients. These values would then be directed to the ANN, which would also access the patient's EHR to determine desired demographic values for the ANN's input vector. The ANN prediction of nonadherence reasons would then be written directly to the EHR so the physician would have this information available should a pharmaceutical treatment plan be diagnosed and enable the physician to direct the patient to various resources to overcome their potential nonadherence. Figure 4 represents the ANN clinical informatics tool relative to other resources.

CONCLUSION

An ANN method for predicting individual patient reasons for nonadherence to treatment has been presented and shown to be able to identify reasons, including multiple reasons, for nonadherence for 64 to 79 percent of the individual patients. Once incorporated into the EHR workflow, physicians will readily have this information available. Physicians or other clinical staff can then use these medical informatics classifications as recommendations on how to plan for ways to overcome their patients' nonadherence

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. ANN clinical informatics tool in healthcare



*The patient survey device may also be used for check in and will then communicate via the office network with the EHR in addition to the INRep ANN

and consequently improve medical outcomes and a healthy lifestyle for these patients. Improving the health outcomes for pharmaceutically treatable conditions may also further promote a healthier population at large through decreased spread of the respective illnesses.

Future research may address how to facilitate incorporation of new features into EHRs and the clinician's natural workflow to enable identification of nonadherence or other dangerous behaviors of patients related to their treatments. Early identification of causes for non-optimal healthcare outcomes will enable education and other efforts, such as automated notifications or reminders, to improve healthcare outcomes.

REFERENCES

Agarwal, S., Yewale, V. N., & Dharmapalan, D. (2015). Antibiotics Use and Misuse in Children: A Knowledge, Attitude and Practice Survey of Parents in India. *Journal of Clinical and Diagnostic Research*, *9*(11), SC21–SC24. PMID:26674397

Akers, R. L. (1977). *Deviant Behavior: A Social Learning Approach*. Belmont, CA: Wadsworth.

Andrajati, R., Tilaqza, A., & Supardi, S. (2016). Factors related to rational antibiotic prescriptions in community health centers in Depok City, Indonesia. *Journal of Infection and Public Health*. Retrieved from http://www.jiph.org/article/S1876-0341(16)00037-X/pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Au, L. E., Harris, S. S., Jacques, P. F., Dwyer, J. T., & Sacheck, J. M. (2014). Adherence to a Vitamin D Supplement Intervention in Urban Schoolchildren. *Journal of the Academy of Nutrition and Dietetics*, *114*(1), 86–90. doi:10.1016/j.jand.2013.06.350 PMID:23999277

Awad, A. I., & Eltayeb, I. B. (2007). Self-medication practices with antibiotics and antimalarials among Sudanese undergraduate university students. *The Annals of Pharmacotherapy*, *41*(7-8), 1249–1255. doi:10.1345/aph.1K068 PMID:17565044

Bourgeois, S., & Yaylacicegi, U. (2010). Electronic Health Records: Improving Patient Safety and Quality of Care in Texas Acute Care Hospitals. *International Journal of Healthcare Information Systems and Informatics*, *5*(3), 1–13. doi:10.4018/jhisi.2010070101

Bowen, D. J., Helmes, A., & Lease, E. (2001). Predicting Compliance: How Are We Doing? In L. Burke & I. Ockene (Eds.), *Compliance in Healthcare and Research*. New York: Wiley.

Buckalew, L. W., & Sallis, R. E. (1986). Patient Compliance and Medication Perception. *Journal of Clinical Psychology*, *42*(1), 49–53. doi:10.1002/1097-4679(198601)42:1<49::AID-JCLP2270420107>3.0.CO;2-F PMID:3950014

Caudill, M. (1987). Neural Networks Primer, Part I. *AI Expert*, *2*(12), 46–52.

Center for Disease Control. (2013). Medication Adherence. Retrieved from https://www.cdc.gov/primarycare/materials/medication/docs/medication-adherence-01ccd.pdf

Cherkassky, V., & Lari-Najafi, H. (1992). Data representation for diagnostic neural networks. *IEEE Expert*, *7*(5), 43–53. doi:10.1109/64.163672

DeMello, J. P., & Deshpande, S. P. (2012). Factors impacting use of information technology by physicians in private practice. *International Journal of Healthcare Information Systems and Informatics*, *7*(2), 17–28. doi:10.4018/jhisi.2012040102

DiMatteo, M. R. (2004). Variations in patients adherence to medical recommendations: A quantitative review of 50 years of research. *Medical Care*, *42*(3), 200–209. doi:10.1097/01.mlr.0000114908.90348.f9 PMID:15076819

DiMatteo, M. R., & Martin, L. R. (2002). *Health Psychology*. Boston: Allyn & Bacon.

Ekwunife, O. I., Udeogaranya, P. O., & Adibe, M. O. (2010). Predictors of self reported adherence to Antihypertensive Medication in a Nigerian population. *Journal of Basic and Clinical Pharmacy*, *1*, 133–138. PMID:24825980

Erah, P. O., & Arute, J. E. (2008). Adherence of HIV/AIDS patients to antiretroviral therapy in a tertiary health facility in Benin City. *African Journal of Pharmacy and Pharmacology*, *2*(7), 145–152.

Eyesan, O. L., & Okuboyejo, S. R. (2013). Design and implementation of a voice-based medical alert system for medication adherence. *Procedia Technology, 2013*(9), 1033-1040.

Farmer, K. C. (1999). Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. *Clinical Therapeutics*, *21*(6), 1074–1090. doi:10.1016/S0149-2918(99)80026-5 PMID:10440628

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gamble, J., Stevenson, M., McClean, E., & Heaney, L. G. (2009). The Prevalence of Nonadherence in Difficult Asthma. *American Journal of Respiratory and Critical Care Medicine*, *180*(9), 817–822. doi:10.1164/rccm.200902-0166OC PMID:19644048

Gibson, T. B., Landrum, M. B., Batata, A., Fendrick, A. M., Wang, S., & Chernew, M. (2011). Regional Variation in Medication Adherence. *Forum Health Econ Policy, 14*(2).

Given, C. W., Given, B. A., & Simoni, L. E. (1978). The association of knowledge and perception of medications with compliance and health states among hypertension patients: A prospective study. *Research in Nursing & Health*, *1*(2), 76–84. doi:10.1002/nur.4770010205 PMID:248835

Gottlieb, H. (2000). Medication Nonadherence: Finding Solutions to a Costly Medical Problem. *Drug Benefit Trends*, *12*(6), 57–62.

Grigoryan, L., Haaijer-Ruskamp, F. M., Burgerhof, J. G., Mechtler, R., Deschepper, R., Tambic-Andrasevic, A., ... Edelsein, H. (2006). Self-medication with antimicrobial drugs in Europe. *Emerging Infectious Diseases*, *12*(3), 452–460. doi:10.3201/eid1203.050992 PMID:16704784

Gupta, A., Woosley, R., Crk, I., & Sarnikar, S. (2007). An Information Technology Architecture for Drug Effectiveness Reporting and Post-Marketing Surveillance. *International Journal of Healthcare Information Systems and Informatics*, *2*(3), 65–80. doi:10.4018/jhisi.2007070106

Hayes, R. B., McDonald, H., Garg, A. X., & Montagne, P. (2002). Interventions for helping patients to follow prescriptions for medications. *The Cochrane Library*, *2*, 1–50.

Helmons, P. J., Suijkerbuijk, B. O., Panday, P. V. N., & Kosterink, J. G. (2015). Drug-drug interaction checking assisted by clinical decision support: A return on investment analysis. *Journal of the American Medical Informatics Association*, *22*(4), 764–772. doi:10.1093/jamia/ocu010 PMID:25670751

Hirth, R. A., Greer, S. L., Albert, J. M., Young, E. W., & Piette, J. D. (2008). Out-of-pocket spending and medication adherence among dialysis patients in twelve countries. *Health Affairs*, *27*(1), 89–102. doi:10.1377/hlthaff.27.1.89 PMID:18180483

Hugtenburg, J. G., Timmers, L., Elders, P. J., Vervloet, M., & van Dijk, L. (2013). Definitions, variants, and causes of nonadherence with medication: A challenge for tailored interventions. *Patient Prefer Adherence*, *7*, 675–682. doi:10.2147/PPA.S29549 PMID:23874088

Ikechuwku, E. O., Obinna, U. P., & Ogochukwu, A. M. (2010). Predictors of Self Reported Adherence to Antihypertensive Medication in a Nigerian Population. *Journal of Basic and Clinical Pharmacy*, *1*(2), 133–138. PMID:24825980

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial Neural Networks: A Tutorial. *Computer*, *29*(3), 31–44. doi:10.1109/2.485891

Kirking, D. M., Lee, J. A., Ellis, J. J., Briesacher, B., & Mckercher, P. L. (2006). Patient-reported underuse of prescription medications: A comparison of nine surveys. *Medical Care Research and Review*, *63*(4), 427–446. doi:10.1177/1077558706288845 PMID:16847072

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kuperman, G. J., Bobb, A., Payne, T. H., Avery, A. J., Gandhi, T. K., Burns, G., ... Bates, D. W. (2007). Medication-related clinical decision support in computerized provider order entry systems: A review. *Journal of the American Medical Informatics Association*, *14*(1), 29–40. doi:10.1197/jamia.M2170 PMID:17068355

Lehane, E., & McCarthy, G. (2007). Intentional and unintentional medication non-adherence: A comprehensive framework for clinical research and practice? A discussion paper. *International Journal of Nursing Studies*, *44*(8), 1468–1477. doi:10.1016/j.ijnurstu.2006.07.010 PMID:16973166

McDonald, H. P., Garg, A. X., & Haynes, R. B. (2002). Interventions to enhance patient adherence to medication prescriptions: Scientific review. *Journal of the American Medical Association*, *288*(22), 2868–2879. doi:10.1001/jama.288.22.2868 PMID:12472329

Okuboyejo, S. (2014). Non-Adherence to Medication in Outpatient Setting in Nigeria: The Effect of Employment Status. *Global Journal of Health Science*, *6*(3), 37–44. doi:10.5539/gjhs.v6n3p37 PMID:24762344

Okuboyejo, S., & Eyesan, O. (2014). mHealth: Using mobile technology to support healthcare. *Online Journal of Public Health Informatics*, *5*(3), 233–242. doi:10.5210/ojphi.v5i3.4865 PMID:24678384

Okuboyejo, S., Omoregbe, N. A., & Mbarika, V. (2012). A Framework for the Design of a Mobile-Based Alert System for Outpatient Adherence in Nigeria. *African Journal of Computing & ICT*, *5*(5), 151–158.

Oyelami, O. M., Okuboyejo, S., & Ebiye, V. (2013). Awareness and usage of Internet-based health information for self-care in Lagos State, Nigeria: Implications for healthcare improvement. *Journal of Health Informatics in Developing Countries*, *7*(2), 165–177.

Pechère, J. C. (2001). Patients interviews and misuse of antibiotics. *Clinical Infectious Diseases*, *33*(Suppl. 3), S170–S173. doi:10.1086/321844 PMID:11524715

Rodvold, D. M., McLeod, D. G., Brandt, J. M., Snow, P. B., & Murphy, G. P. (2001). Introduction to artificial neural networks for physicians: Taking the lid off the black box. *The Prostate*, *46*(1), 39–44. doi:10.1002/1097-0045(200101)46:1<39::AID-PROS1006>3.0.CO;2-M PMID:11170130

Smith, M. (1993). *Neural Networks for Statistical Modeling*. New York: Van Nostrand Reinhold.

Urquhart, E. (2005). Challenging perceptions of long-acting injectable antipsychotic medications. *Mental Health Practice*, *8*(10), 14–18. doi:10.7748/mhp2005.07.8.10.14.c1872

Uzochukwu, B. S. C., Onwujekwe, O. E., Onoka, A. C., Okoli, C., Uguru, N. P., & Chukwuogo, O. I. (2009). Determinants of Non-Adherence to Subsidized Anti-Retroviral Treatment in Southeast Nigeria. *Health Policy and Planning*, *24*(3), 189–196. doi:10.1093/heapol/czp006 PMID:19276155

Walczak, S., & Cerpa, N. (1999). Heuristic Principles for the Design of Artificial Neural Networks. *Information and Software Technology*, *41*(2), 107–117. doi:10.1016/S0950-5849(98)00116-5

Wapner, J. (2008). Doctors Without Orders The Mystery of Patients Who Fail to Follow Prescriptions. *Medical Examiner*. Retrieved from http://www.slate.com/articles/health_and_science/medical_examiner/2008/03/doctors_without_orders.html

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zhang, G. P. (2007). Avoiding pitfalls in neural network research. *IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews*, *37*(1), 3–16. doi:10.1109/TSMCC.2006.876059

Zhu, X., Pan, H., Yang, Z., Cui, B., Zhang, D., & Ba-Thein, W. (2016). Self-medication practices with antibiotics among Chinese university students. *Public Health*, *130*, 78–83. doi:10.1016/j.puhe.2015.04.005 PMID:26008209

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 12(1); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 1-13; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 28
Applying CI in Biology through PSO

Rojalina Priyadarshini
C. V. Raman College of Engineering, India

Brojo Kishore Mishra
C. V. Raman College of Engineering, India

Nilamadhab Dash
C. V. Raman College of Engineering, India

Rachita Misra
C. V. Raman College of Engineering, India

ABSTRACT

Conventional computing methods face challenges dealing with real world problems, which are characterised by noisy or incomplete data. To find solutions for such problems, natural systems have evolved over the years and on analysis it has been found these contain many simple elements when working together to solve real life complex problems. Swarm Intelligence (SI) is one of the techniques which is inspired by nature and is a population based algorithm motivated by the collective behaviour of a group of social insects. Particle swarm optimization (PSO) is one of the techniques belonging to this group, used to solve some optimization problems. This chapter will discuss some of the problems existing in computational biology, their contemporary solution methods followed by the use of PSO to address those problems. Along with this several applications of PSO are discussed in few of the relevant fields are discussed having some future research directions on this field.

INTRODUCTION

A massive growth is seen in gathering biological information by scientific communities in these years. This information comes in the form of genomes, protein sequences and gene expression data and so on which causes a need for and efficient computational tools to store, analyze and construe these data. The term bioinformatics and CB always used interchangeably in the field of computer science, which literally means the science of informatics as applied to biological research. Informatics on the other hand is the management and analysis of data using various advanced computing techniques. Hence, in other words, bioinformatics can be described as the application of computational methods to make biological discoveries and explorations (Clote & Backofen, 2000). It can also be defined as the appliance of

DOI: 10.4018/978-1-7998-1204-3.ch028

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

computer technology to the management of biological information, encircling a study of the inherent genetic information, underlying molecular structure, resulting biochemical functions, and the exhibited phenotypic properties (Sushmita Mitra, Yoichi Hayashi,(2006)). The final goal and attempt of this area is to create new insights into the science of life as well as developing a global perspective (Das, Abraham & Konar, 2008). Three major objectives of bioinformatics can be put forward as to develop:

- Algorithms and mathematical models for searching the relationships among the members of huge biological dataset.
- Analyze and interpret the variety of heterogeneous data including nucleotide and amino acid sequences, protein domains and protein structures.
- Tools that enable proficient and effective storage, retrieval and management of high-volume biological databases.

In recent times, large number of biologically motivated algorithms have been developed and are being used for handling many complex problems. For instance, neural computing (Haykin, 1999) attempts to imitate the biological nervous systems of the living creatures inspired by the amount of parallel and distributed processing done inside the brain of a human. Genetic algorithm (GA) is inspired by the Darwinian evolutionary process through cross-over and mutation of biological chromosomes. They have successfully been used in many bioinformatics tasks that need intelligent search, optimization and machine learning approaches. Sushmita Mitra and Yoichi Hayashi, (2006) provide a comprehensive survey of the research in this direction. For the past few years there has been a slow but steady increase of research papers reporting the success Computational Intelligence (CI) based searching, classification and clustering techniques applied to the field of **CB** (Das, Panigrahi & Pattnaik, 2009). CI is constituted with the study of a set of nature-inspired computational methodologies, techniques and approaches to address complex real-world problems to which traditional approaches are ineffective or infeasible. While traditional models often fail to handle uncertainty, noise and the presence of an ever-changing context, CI provides solutions for complicated real problems. It primarily includes artificial neural networks (ANN), evolutionary computation (EC) and fuzzy logic (FL). In addition, CI also embraces biologically inspired algorithms such as SI, artificial immune systems, which can be seen as a part of EC.

This chapter presents some fundamental problems present in CB, the existing and traditional solution methods to answer these problems, the issues and challenges associated with these. Some of the contemporary problems solving approaches based on CI are then discussed.

BACKGROUND

To understand bioinformatics in any meaningful way, it is necessary for a computer scientist to understand some basic biology. In this chapter a short and basic introduction to the fundamentals and relevant to computational thing is given which are illustrated below.

Genetic Material: An Overview

Biology at the microscopic level began in 1665 when, Robert Hooke discovered that organisms are composed of individual compartments called cells. A cell's traits were inherent in its genetic information,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

which was passed to its offspring, and the same was organized into genes that resided on chromosomes. Later on proteins and DNA (Dioxyribo Nucleic Acids) were discovered in 1869 by Johann Friedrich Miescher which consist of four types of bases: Adenine (A), Thymine (T), Guanine (G), and Cytosine (C). Later biologists discovered that not all Ribo Nucleic Acids (RNAs) are destined to serve as templates for building proteins. Some RNAs (like transfer RNA described below) play a different role.

DNA: TAC CGC GGC TAT TAC TGC CAG GAA GGA ACT

RNA: AUG GCG CCG AUA AUG ACG UC CUU CCU UGA

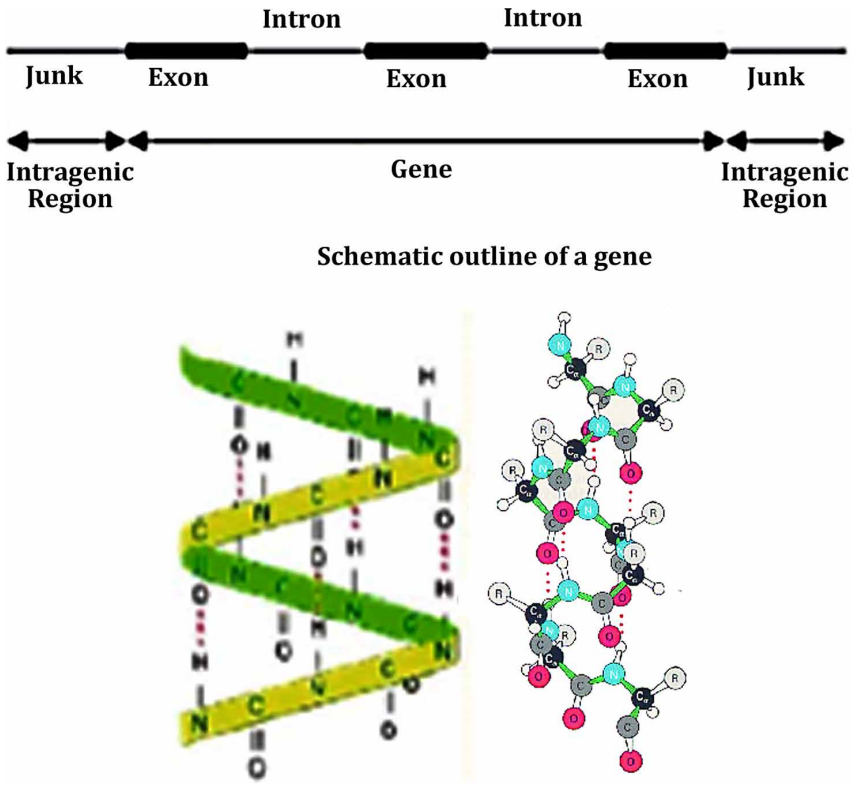
Protein: Met Ala Pro Ile Met Thr Val Leu Pro Stop

The process of turning an mRNA into a protein is called translation, since it translates information from the RNA (written in a four-letter alphabet) into the protein (written in 20-letter alphabet). All proteins, including the ones necessary for this process, are produced by this process. This flow of information is emphatically referred to as the central dogma in molecular biology.

DNA → transcription → RNA → translation → protein,

Figure 1 describes the analysis of DNA.

Figure 1. DNA analysis



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MAIN FOCUS OF THE CHAPTER

In this section, we provide a basic understanding of the fundamentals problems of computational biology like protein structure, folding, DNA microarray data interpretation, inferring the gene regulatory networks that are relevant to this article.

Protein Structure Prediction

Proteins are built up by polypeptide chains of amino acids, which consist of DNA as the building block. An amino acid is an organic molecule made up of an amine (NH) and a carboxylic (CO) acid group (backbone), collectively bonded with a side-chain (hydrogen atom and residue R) that creates separation between them. Proteins are polypeptides, created in cells as a linear chain of amino acids. Chemical properties that differentiate the 20 individually different amino acids cause the protein chains to fold up into specific3-D structures that define their particular functions in the cell. Given the primary structure of a protein, in terms of a linear sequence of amino acids, folding tries to predict its stable 3-D structure. All these interactions governed by the laws of physics and chemistry to predict 3-D positions of different atoms in the protein molecule. Protein folding is a reaction involving other interacting amino acids and water molecules. The two-dimensional (2-D) secondary structure can involve an α-helix (with the CO group of the ith residue hydrogen (H) bonded to the NH group of the $(i+4)^{th}$ one) or a β-sheet (corrugated or hairpin structure) formed by the H-bonds between the amino acids. The parts of the protein that are not characterized by any regular H-bonding patterns are called random coils or turns. The tertiary structure refers to the 3-D conformation of the protein.

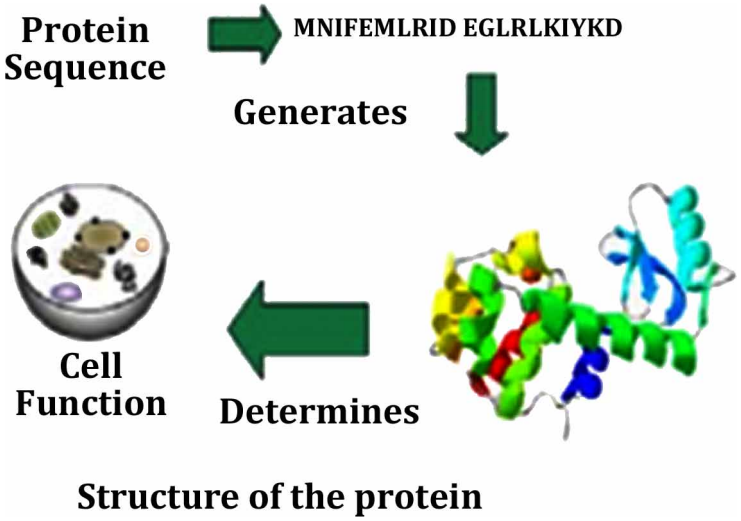
The total number of experimentally determined structure is less compare to actually known protein sequences (Schulze-Kremer, 2000). The difficult determination of these structures, using experimental methods such as X-ray or nuclear magnetic resonance (NMR), contributes to increase this gap between sequence and protein structures. Therefore, it is necessary the use of computational methods which predict 3D protein structures in a cheaper and faster way. According to biological approximation the methods for PSP falls into three groups: (1). Homology based Comparative modeling, (2) *ab* initio methods and (3). Threading (Priyadarshini, Dash, & Rout, 2012). The comparative model uses the predicted structures from the Protein data bank. These methods are based on the hypothesis that similar protein sequences determine similar 3D structures. Figure 2 shows the relation between protein sequence and their structure.

The problem of protein structure prediction has two dimensions:

1. **Secondary Structure:** A step on the way to a prediction of the full 3-D structure of protein is predicting the local conformation of the polypeptide chain, called the secondary structure. The whole framework was pioneered by Chou and Fasmann. They used a statistical method, with the likelihood of each amino acid being one of the three (alpha, beta, coil) secondary structures estimated from known proteins. The data consist of proteins obtained from the PDB. A fixed size window constitutes the input to the machine learning model. The machine classifier predicts the secondary structure corresponding to the centrally located amino-acid of the sequence within the window. The contextual information about the rest of the sequence in the window is also considered according to the model of the machine.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Protein Sequence-Structure-Function Relationship



2. **Tertiary Structure and Folding:** Protein structure comparison is often used to identify set of residue equivalencies between proteins based on their 3-D coordinates, and has a wide impact on the understanding of protein sequence, structure, function, and evolution. The determination of an optimal 3-D conformation of a protein corresponds to folding, and has manifold implications to drug design. An active site structure determines the functionality of a protein. A ligand (enzyme or drug) docks into an active site of a protein. Many automated docking approaches have been developed, and can be categorized as: 1) rigid docking: both ligand and protein are rigid; 2) flexible-ligand docking: ligand flexible and protein rigid; and 3) flexible-protein docking: both ligand and protein are flexible (only a limited model of protein variation allowed, such as side-chain flexibility or small motions of loops in the binding site).

Primary Genome Sequence Analysis

The nucleus of a cell carries chromosomes that are built upon the double helical DNA molecules. In the human case the DNA genome is arranged into 24 dissimilar chromosomes. Each chromosome contains a number of genes, which are the fundamental physical and functional units of heredity. But, genes encompass only about 2% of the human genome; the remainder consists of non-coding regions, whose functions may embrace providing chromosomal structural integrity and regulating where, when, and in what amount and quantity proteins are built. DNA is transcribed to generate messenger (m)-RNA, which is then translated to make protein. The m-RNA is single stranded and has a ribose sugar molecule. A gene contains "Promoter" and "Termination" sites, which are responsible for the initiation and termination of transcription. Translation consists of mapping from triplets (codons) of four bases to the 20 amino acids building block of proteins. Genes are typically organized as exons (coding regions) and introns (noncoding regions). Hence, the main task of gene identification, from the primary genomic sequence, involves coding region recognition and splice junction4 detection. Sequence data are typically dynamic and order-dependent. A protein sequence motif is a signature or consensus pattern that is embedded

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

within sequences of the same protein family. Identification of the motif leads to classification of an unknown sequence into a protein family for further biological analysis. Available protein motif databases include PROSITE5 and PFAM.

Microarrays Data Analysis

A microarray is a glass slide on to which DNA molecules are statically set in an orderly manner at precise and definite positions called spots (or features). A microarray may contain thousands of spots and each spot may contain a few million copies of identical DNA molecules that uniquely correspond to a gene. This measures the relative m-RNA profusion between two samples, which are labelled with different flourescent dyes, viz., red and green. The m-RNA binds (hybridizes) with the Codon DNA(cDNA) probes on the array. The relative abundance of a spot or gene is measured as the logarithmic ratio between the intensities of the dyes, and constitutes the gene expression data. Gene expression levels can be determined for samples taken:

1. At multiple time instants of a biological process (different phases of cell division) or
2. Under various conditions (e.g., tumor samples with different histopathological diagnosis).

Each gene corresponds to a high-dimensional vector of its expression profile. The data contain a high level of noise due to experimental procedures. Moreover, the expression values of single genes demonstrate large biological variance within tissue samples from the same class. A major cause of coexpression of genes is their sharing of the regulation mechanism (coregulation) at the sequence level. Clustering of coexpressed genes, into biologically meaningful groups, helps in inferring the biological role of an unknown gene that is coexpressed with a known gene(s). Cluster validation is essential, from both the biological and statistical perspectives, in order to biologically validate and objectively compare the results generated by different clustering algorithms.

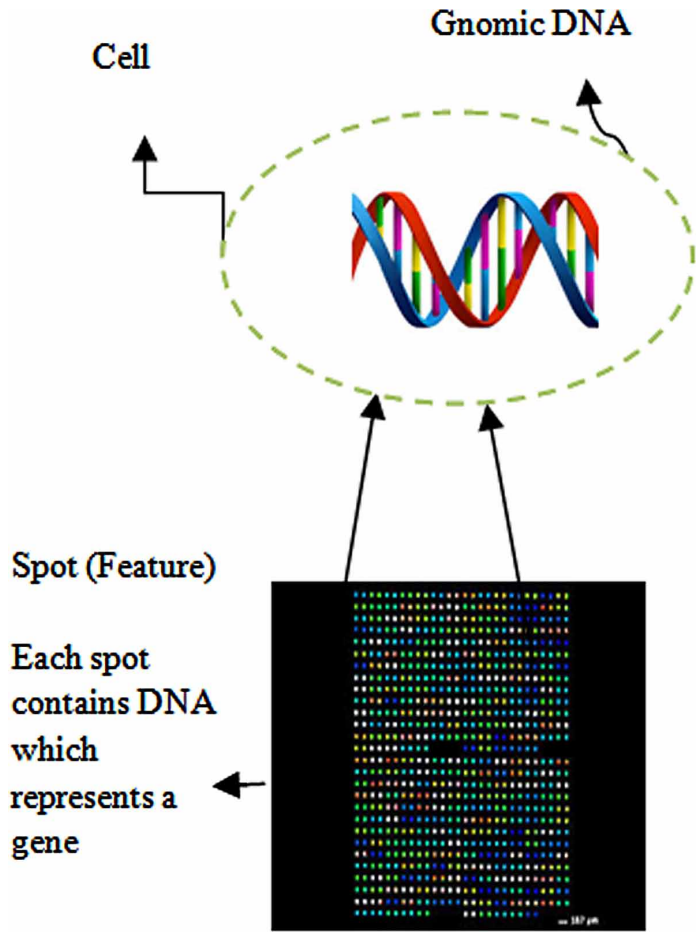
In Figure 3, a microarray may contain thousands of spots. Each spot contains many copies of the same DNA sequence that uniquely represents a gene from an organism. Spots are arranged in an orderly fashion into Pen groups. Figure 4 represents schematic of the experimental protocol to study differential expression of genes. The organism is grown in two different conditions (a reference condition and a test condition). RNA is extracted from the two cells, and is labelled with different dyes (red and green) during the synthesis of cDNA by reverse transcriptase. Following this step, cDNA is hybridized onto the microarray slide, where each cDNA molecule representing a gene will bind to the spot containing its complementary DNA sequence. The microarray slide is then excited with a laser at suitable wavelengths to detect the red and green dyes (Das et al. 2006).

Understanding Gene Regulatory Network

A gene regulatory network or genetic regulatory network (GRN) is a group of related regulators intermingle with each other and with other external substances in the cell which are responsible to manage the gene expression levels of mRNA and proteins. The regulator can be RNA, protein and their complex. The interaction can be direct or indirect (through their transcribed RNA or translated protein). In this processes they may generate mass, energy, information transfer, and cell-fate specification. The whole process of a cell during this is integrated through a complex network of cellular constituents and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

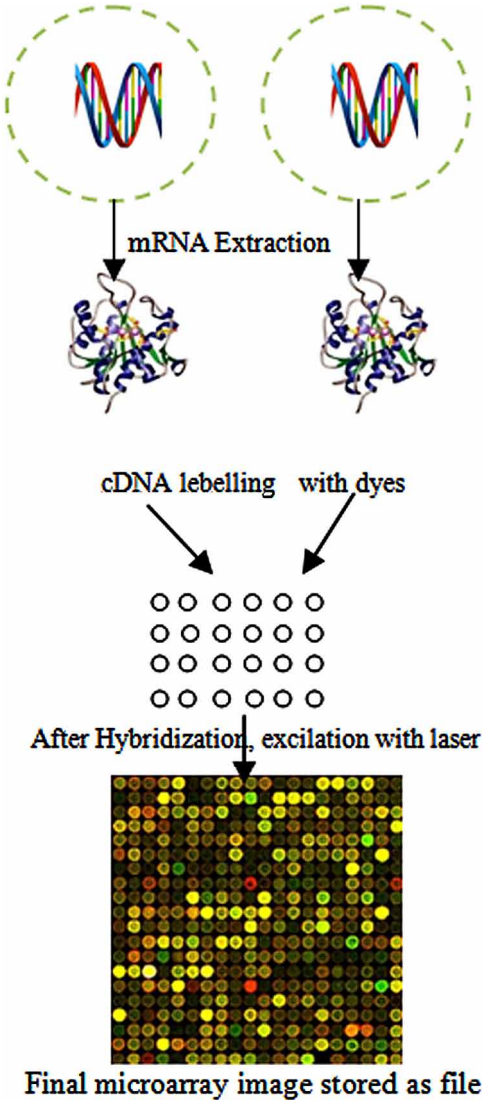
Figure 3. Microarray expressions



reactions. Such a metabolic network consists of nodes, i.e., substrates (genes or proteins), that are interconnected through links. The degree of interconnectivity of the network may be characterized by its diameter, which is the shortest biochemical pathway averaged over all pairs of substrates. The topology of a network reflects a long evolutionary process molded for a robust response toward internal defects and environmental fluctuations. Despite significant variation of individual constituents and pathways, metabolic networks have the same topological scaling properties and exhibit striking similarities to the inherent organization of complex, robust non-biological systems. It is a knowledge base for systematic analysis of gene functions in terms of the networks of genes and molecules. The data objects are represented as graphs, and various computational methods are developed to detect graph features that can be related to biological functions. For example, it can: 1) reconstruct biochemical pathways from the complete genome sequence; 2) predict gene regulatory networks from gene expression profiles, obtained by microarray experiments; and 3) Single-stranded DNA that is complementary to mRNA or DNA that has been synthesized from messenger RNA by the enzyme reverse transcriptase. Predicting the gene expression levels is a main task associated with these networks. Along with the conventional approaches like directed graph, Boolean network, Bayesian network etc. some machine learning approaches is used to model these networks (Chan, Havukkala, Jain, Hu, & Kasabov, 2008). So that they can be interpreted

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Experimental gene expression data



well, this in turn predicts the expression level more accurately. If the level of accuracy percentage is more, then it provides a better way to explore how drugs affect a system of genes as well as for finding which genes are interrelated in a particular process.

Sequence Alignment

It is the way of arranging biological sequences such as DNA, RNA, or protein sequence to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences. Aligned sequences of nucleotide or amino acid residues are typically represented as rows within a matrix. Gaps are inserted between the residues so that identical or similar characters are aligned in successive columns. These techniques are grouped into two categories a) pair-wise sequence alignment

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

b) Multiple sequence alignment. If alignment is done on two sequences, it is said to be pair wise else it is said to be multiple alignment. Either the whole input sequence is aligned from the beginning to the end of the sequence; otherwise some of the conserved regions are only aligned with the given sequence. In the first case it is said to be the global alignment and the latter case is said to be local alignment.

CI METHODS FOR SOLVING COMPUTATIONAL BIOLOGY PROBLEMS

Several CI methods like ANN, fuzzy logic, GA, SVM are being used for solving different kinds of bio-informatics problems, which are discussed in this section.

Methods for Solving Protein Structure Prediction

The main SC paradigms for the application of PSP are ANN, EC and SVMs (Márquez-Chamorroa, Asencio-Cortes, Santiesteban-Toca & Aguilar-Ruiz, 2015).

Neural Network Methods

The authors in Casadio, Compiani, Fariselli, and Vivarelli (1995) present a feed-forward neural network which is trained with matching sets of amino acid sequences and two different types of structural information. The corresponding secondary structure and the contact map of known proteins. The input of the network consists of a window of 61 residues. The hidden layer has 300 neurons and the output layer 30 neurons used to predict contacts between the amino acid that occupies the central positions in the window and the rest of the residues of the window. Three additional neurons are used to predict the tertiary motifs. The method proposed in Cho and Ryu (2002) is based on two input neurons which use input vectors with information about the residue–residue contact and its environment, the number of amino acids and the separation between amino acids. In addition, several variables are added, such as the hydrophobicity of the environment, as well as evolutionary information.

A supervised feed-forward network which is Radial basis function (RBF) network, is being used (Casadio, Compiani, Fariselli & Vivarelli, 1995) to optimally predict the free energy contributions of proteins due to hydrogen bonds, hydrophobic interactions, and the unfolded state, with simple input measures.

Support Vector Machines

SVM techniques performs classification tasks building a hyper-plane in a multidimensional space, trying to maximize the margin between each different classes. The function that performs the transformation of the space is called kernel function. SVMs are used as a machine learning tool to predict tertiary structure from the primary sequence. The input to the classifier are the result of multiple sequence alignment (MSA) (Cheng & Baldi, 2007) data of similar proteins after transformation to a specific input domain, amino acid physic chemical properties or contact potentials, local window features of the sequences.

As limitations of these types of methods, it can be cited that the kernel models over fits the model selection criterion, the difficulty in the selection of the optimal kernel function parameters and the algorithmic complexity and extensive memory requirements in large-scale tasks.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Evolutionary Computation (EC)

Methods based on EC may use various possible representations of a protein structure: dihedral or torsion angles and lattice models. According to the representation models the evolutionary models may differ. Some methods use dihedral or torsion angles some use 3- torsion angles without using the mutation operator (Cui, Chen & Hung, 1998; Schulze-Kremer, 2000). The fitness function consists of hydrophobic interactions and Vander Waals contact measures. A force field model, that represents chemical reactions and physical forces that occur in a protein is developed for GA. Mainly the fitness function used here are for minimizing the potential energy of the hydrogen bonds of the amino acids which makes a 3D structure of a protein as stable.

Methods for Genome Sequence Analysis

Neural Networks

ANN is used for predicting both types of genes, that is, protein coding and RNA coding (Goel, Singh & Aseri, 2013). One of the earliest attempts to ANN network for gene prediction is made in 1991, GRAIL (gene recognition and analysis internet link). A multilayer feed-forward neural network is applied in GRAIL-I, that accepts input from seven statistical measures taken on a 99-base window. A similar system to GRAIL came in 1993. A program named Gene Parser, which is employed to predict protein genes in genomic DNA sequences. This program uses the content statistics as well as site statistics to predict exons, introns, and their boundaries. Both single and multilayer neural networks are used here to unite information from these statistics and database search information. Here a recursive optimization procedure based on dynamic programming is used to find the most probable combinations of exon and introns. Many gene prediction techniques utilize homology information, which helps in improving the prediction results. In the latest version of Grail, homology information has been incorporated. The new system resulted in improved performance and named as GrailExp (Vimaladevi & Kalaavathi, 2014). A computational gene prediction system GIN (gene identification using neural nets and homology information) was developed in 1998 to avoid false positive predictions. The technique combines homology information from protein and expressed sequence tag databases into back-propagation neural network (Márquez-Chamorroa et al. 2015). The program can recognize multiple genes within genomic DNA. GIN performs better than other methods (e.g., GeneID or GeneParser3) that make use of homology information to predict genes. The performance of the system is better than GENSCAN in gene level accuracy. This technique does not work well in the absence of homology information.

Methods for Microarray Data Analysis

Microarray data analysis domain is the management of complicated experimental data for application of computational methods in biology. Complicated experimental data causes two types of problems. The first one is data need to be pre-processed. Pre-processing is the process by which the data needs to be modified which would be suitably used by computational intelligence methods. The second is the analysis of the specific data which depends on what we intend to search for. Most microarray data analysis packages implement five distinct steps in microarray data analysis (Dembele & Kastner, 2003; Babu, 2004):

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. Data pre-processing (intra-chip and inter-chip normalisation).
2. Gene Selection (identifying differentially expressed genes).
3. Clustering (identifying common expression patterns – co-expression analysis).
4. Functional Enrichment/Biological Pathway analysis (identifying the biological significance of the selected genes).
5. Classification (developing a classification system for unclassified samples).
    a. Fuzzy Systems

Fuzzy c-means is a popular fuzzy based algorithm used for clustering microarray overlapping data. Different features of the functions regulations can be explored by fuzzy clustering which enables genes to belong to multiple groups at the same time. Fuzzy c-means algorithm has been applied to cluster microarray data (Sarhan, 2009). The value of the fuzzifier is appropriately chosen and fine-tuned for selection of genes, based on resultant distribution of distances between them. The clusters are formed by tight association and coupling among the genes.

ANNs

The two major mining tasks, modelled here, are clustering and classification. While unsupervised learning is self-organized, supervised learning helps incorporate known biological functions of genes into the knowledge discovery process of gene expression pattern analysis for gene discovery and prediction:

1. **Clustering:** Kohonen's Self Organising Maps (SOM) are used to make clusters of gene expression data (Dembele &. Kastner, 2003). It provides a good visualization as well as generates a robust and accurate clustering of large and noisy data. SOMs require a selected node in the gene expression space (along with its neighbours) to be rotated in the direction of a selected gene expression profile (pattern). However, the pre definition of a 2-D topology of nodes can often be a problem considering its biological relevance.
2. **Classification:** A wide range of supervised learning algorithms have been applied to microarray data for sample classification. Techniques ranging from ANN to SVM, from K Nearest Neighbour (KNN) to Naïve Bayesian Classifiers (NBC) have been applied to microarray data classification and their adequacy assessed.

Classification of acute leukemia has been made by combining a pair of classifiers trained with mutually exclusive features (Cho & Ryu, 2002). A neural network combines the outputs of the multiple classifiers. Feature selection with non-overlapping correlation encourages the classifier ensemble to learn different aspects of the training data in a wide solution space. Ahmad M. Sarhan (2009) has presented a stomach cancer detection system based on ANN, and the Discrete Cosine Transform (DCT). DCT has been used to extract classification features from stomach micro arrays in this system. The extracted features from the DCT coefficients are being applied to an ANN for classification (tumor or non-tumor). The micro array images used in this study have been obtained from the Stanford Medical Database (SMD). Bharathi and Natarajan (2010) have tried to fetch the minimum set of genes that can ensure accurate classification of cancer to a higher rate from micro array data. They used supervised machine learning algorithms. Bo Li, Chun-Hou Zheng, De-Shuang Huang, Lei Zhang and Kyungsook Han (2010) have discussed that the gene expression data collected from DNA micro array are characterized by a large amount of genes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

but with only a small amount of observations (experiments). They have proposed a manifold learning method which can map the gene expression data to a low dimensional space, and then tried to discover the internal structure of the features which could classify the micro array data with more accuracy. The proposed algorithm could able to bring the gene expression data into a subspace with high intra-class compactness and inter-class separability. Their experimental results on six DNA micro array datasets have shown that the used method is effective for discriminant feature extraction and gene expression data classification (Vaishali & vinayababu, 2011).

SOLUTIONS AND RECOMMENDATIONS

SI is a discipline which deals with natural and artificial systems composed of many individual components which coordinate among themselves using decentralized control and self-organization (Blum & Merkle, 2008). This discipline focuses on the collective behaviour of individual objects with local interactions of the individuals with each other and with their environment which comprises of Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), and Ant Colony Optimization (ACO).

A swarm is a cluster of same type of, simple agents which interacts among themselves locally, and their environment. They do not have central control over themselves. Swarm-based algorithms have recently emerged as a new class of nature-inspired, population-based algorithms that are capable of providing low cost, fast, and robust solutions to several complex real life problems. SI can be defined as a relatively new branch of Artificial Intelligence that is used to model the collective behaviour of social swarms in nature, such as ant colonies, honey bees, and bird flocks. Although these agents (insects or swarm individuals) are having limited capabilities on their own, they are interacting together with certain behavioural patterns to cooperatively achieve the goal towards the tasks necessary for their survival (Belal, Gaber, El-Sayed & Almojel, 2006; Dorigo, Bonabeau & Theraulaz, 2000). There can be direct or indirect social interactions among swarm objects. The pheromone trail secretion can be taken as the example of indirect contact where on a change by the individual in the present environment, other objects respond to this. The indirect interaction which involves the environment is referred to as stigmergy, means communication with the help of the environment.

Characteristic of a Swarm Intelligence System

The typical swarm intelligence system has the following characteristics:

- It is composed of more than one individual objects, each called as an agent or a particle;
- The particles or agents are relatively homogeneous (i.e., they are either all identical or they belong to a particular and specific category);
- The rules through which the interactions among the individuals happens are based on local information that the individuals exchange either directly or indirectly through the surroundings (stigmergy);
- The interactions of individuals with each other and with their environment is represented as the overall collective behaviour or group behaviour of the system results from their interaction, which is termed as, the group behaviour self-organizes. (Dorigo, 1992; Karaboga, 2005; Karaboga & Basturk, 2007).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The collective and social behaviour of living creatures inspired researchers to commence the study of SI (Lim, Jain, & Dehuri, 2009). The phrase Swarm Intelligence was coined by Beny and Wang in late 1980s (Couzin ID, Krause, James, Ruxton & Franks, 2002) in the context of cellular robotics. In spite of having no centralized control to guard how individual agents should behave, they are controlled globally by local interactions between themselves which lead to the emergence of global behavior (Das, Panigrahi & Pattnaik; Panigrahi, Shi & Lim, 2011). These examples are found in our surrounding such as fish-schools and bird flocks. may change shape and direction, they come into sight to move as a single coherent entity. The main properties of collective behavior are shown in Figure 5:

- **Homogeneity**: Every bird in flock has the same behavioural model. Without any specific leader flock moves, even though momentary leaders seem to emerge.
- **Locality**: Its nearest flock-mates are only influenced by the motion of each bird. Vision is considered to be the most important senses for flock organization.
- **Collision Avoidance**: They avoid colliding with nearby flock mates.
- **Velocity Matching**: They try to match with the velocity of nearby flock mates.
- **Flock Centering**: All of them in the cluster attempt to reside in close proximity with the nearby flock mates. Each of them attempts to maintain a minimum distance between them sat all times. (Couzin ID et.al, 2002) identified four collective dynamical behaviors as illustrated in Figure 4.
- **Swarm**: An aggregate with cohesion, but a low level of polarization (parallel alignment) among members.
- **Torus**: Individuals perpetually rotate around an empty core (milling). The direction of rotation is random.
- **Dynamic Parallel Group**: The individuals are polarized and move as a coherent group, but individuals can move throughout the group and density and group form can fluctuate.
- **Highly Parallel Group**: Much more static in terms of exchange of spatial positions within the group than the dynamic parallel group and the variation in density and form is minimal (Panigrahi & Pattnaik, 2009).

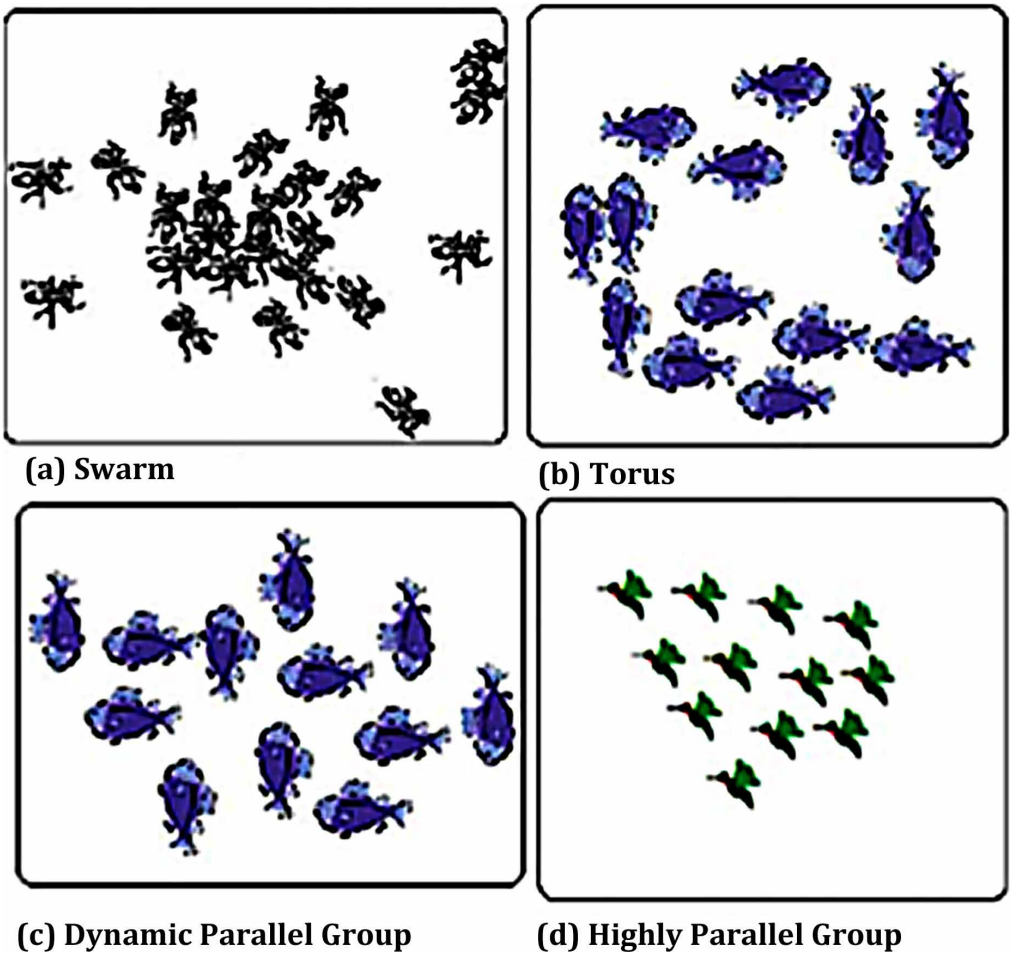
The SI algorithms are mainly stochastic search and optimization techniques, which are followed by the principles of collective behaviour and self-organization of insect swarms. Coincidentally many jobs in bioinformatics entail optimization of different parameters (like energy, alignment score, overlap strength and so on. For finding the solutions of bioinformatics problems always we go for some in-exact but robust, fast and nearly optimal solutions, for which SI based PSO algorithm are suitable (Engelbrecht, 2002).

PSO: An Insight

Particle swarm optimization (PSO) is a population based stochastic search and optimization technique, which was introduced by Kennedy and Eberhart in (1995). It is a multi-objective optimization method to find optimal solution to the problems having multiple objectives. It does not require any gradient information of the function to be optimized uses only primitive mathematical operators and is conceptually very simple. Particles are conceptual entities which fly through the multi-dimensional search space. At any particular instant, each particle has a position and a velocity. The position vector of a particle with respect to the origin of the search space represents a trial solution of the search problem. At the beginning a population of particles is initialized with random positions marked by vectors Xi and random veloci-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Different models of collective behavior (adapted from 8)



ties V_i. The population of such particles is called a 'swarm' S. A neighbourhood relation N is defined in the swarm. N determines for any two particles Zi and Zj whether they are neighbours or not. Thus for any particle Z, a neighbourhood can be assigned as N(Z), containing all the neighbours of that particle. The goal is to reach to the global optimum of some multidimensional and possibly nonlinear function or system (Hsieh, 2014). Initially particles are randomly distributed over the search space. So each particle gets a virtual position that represents a possible potential solution to optimization problem. It is an iterative process where in each iteration every particle moves to a new position, navigating through the entire search space. Each particle keeps track of its position in the search space and its best solution so far achieved. The personal best value is called as pBest and the ultimate goal is to find the global best called as gBest. The standard PSO algorithm broadly consists of the following computational steps

Initialize particles with random positions and velocities;
For each value of k [where 'k' represents the number of particles]
do
Evaluate fitness of each particle's position (p)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

If fitness(p) better than fitness(pbest) then pbest= p
Set best of pBests as gBest
Update particles velocity and position
End of for
Stop: giving gBest, optimal solution. (www.swarmintelligence.org(retrived on 20-07-02015))
Here, a particle refers to a potential solution to a problem in d- dimensional design space with k particles.
Each particle is characterized by
Position vector….. xi(t) and Velocity vector…...vi(t)

Each particle has individual knowledge pbest, its own best-so-far position, Social knowledge gbest, pbest of its best neighbour

The equations for velocity and position updates are given below

$$v(t+1)=(w * v(t))+(c1 * r1 *(p(t) – x(t))+(c2*r2*(g(t) – x(t)) \quad (1)$$

$$x(t+1) = x(t) + v(t+1) \quad (2)$$

The first equation updates a particle's velocity. The term v (t+1) is the velocity at time t+1. The new velocity depends on three terms. The first term is w * v (t). The w factor is called the inertia weight and is just a constant between 0 to 1. Here the value of w is taken as 0.73, and v (t) is the current velocity at time t.

The second term is c1 * r1 * (p (t) – x (t)). The c1 factor is a constant called the cognitive (or personal or local) weight. The r1 factor is a random variable in the range [0, 1) which is greater than or equal to 0 and strictly less than 1. The p (t) vector value is the particle's best position found so far. The x(t) vector value is the particle's current position.

The third term in the velocity update equation is (c2 * r2 * (g (t) – x (t)). The c2 factor is a constant called the social, or global, weight. The r2 factor is a random variable in the range [0, 1). The g (t) vector value is the best known position found by any particle in the swarm so far. Once the new velocity, v (t+1) has been determined, it is used to compute the new particle position x (t+1).

The updates in PSO can be either synchronous or asynchronous (Aziz, Mubin, Saberi Mohamad & Aziz2, 2014).

Types of PSO

Binary PSO

PSO was originally developed to optimize continuous valued parameters. Kennedy and Eberhart proposed a binary PSO to optimize binary-valued parameters. Here position vectors are binary vectors, and the velocity vectors are still floating-point vectors.

However, velocities are used to determine the probability that an element of the position vector is bit 0 or bit 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Angle Modulated PSO

It is developed by Pampara, Engelbrecht and Franken to optimize binary valued parameters by evolving a bit string generating function,

$$g(x)=sin\,(2\pi\,(x-a)\times b\times cos(2\,\pi\,(x-a)\times c))+d$$

The task is then to find values for a,b,c and d, where these values are floating-points.
Some Advantages of the basic PSO are:

- It can be applied into both scientific research and engineering use.
- PSO have no overlapping and mutation calculation. The search can be carried out by the speed of the particle.
- PSO adopts the real number code, and it is decided directly by the solution. The number of the dimension is equal to the constant of the solution (Qinghai Bai (2010.)

Disadvantages of the basic particle swarm optimization algorithm:

- The method easily suffers from the partial optimism, which causes the less exact at the regulation of its speed and the direction.
- The method cannot work out the problems of scattering and optimization (Yonggang, Fengjie & Jigui, 2006 (In Chinese).
- The method cannot work out the problems of non-coordinate system, such as the solution to the energy field and the moving rules of the particles in the energy field (Bai, 2010).

Use of PSO in Computational Biology Problems

In this section the solutions of the some of the problems in computational biology are discussed using PSO for some of the reason.

PSO for Clustering Gene Expression Data

The standard clustering algorithms such as hierarchical clustering, principle component analysis (PCA), GA and ANNs have been used to cluster gene expression data. But in 2003, Xiao et al. used a new approach based on the hybridization of PSO and the Self Organizing Maps (SOM) for making clusters on them. Authors achieved good and encouraging results by using hybrid SOM-PSO algorithm over the gene expression data of Yeast and Rat Hepatocytes (Xiao, Dow, Eberhart, Miled & Oppelt, 2003).

Jun Sun et al. in 2012 has used the quantum behaved-PSO for gene expression data analysis with clustering methods. The problem of clustering gene expression data consists of N genes and their corresponding N patterns. Each pattern is a D-dimensional vector storing the expression levels of the genes at each of the D time points. The objective of applying MEQPSO clustering algorithm is to divide a gene expression data set into user-defined K groups by taking Euclidean distance as similarity measure, such that this partition minimizes the Total Within-Cluster Variation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

PSO for Protein Structure Prediction

Nashat Mansur et al. in 2012 present a particle swarm optimization (PSO) based algorithm for predicting protein structures in the 3D hydrophobic polar model. Starting from a small set of candidate solutions, their algorithm claimed to efficiently explore the search space and returns 3D protein structures with minimal energy. A particle is a candidate solution given by an array of length n (with index 0, 1, . . ., n-1), where n is the number of amino acids in the respective protein. Each element in the array represents the position X_d of the corresponding amino acid 'd' with respect to the preceding one and its value can be one of six characters {b, f, u, d, l, r} indicating six directions {backward, forward, up, down, left, right}. Initially, the swarm is populated with a set of N candidate solutions which are randomly generated. All the velocities are initially set to 0. A particle is invalid if it experiences collision. Collision occurs if two or more amino acids lie at the same point on the cubic 3D lattice. Invalid particles are not accepted in their algorithm and are repaired using a backtracking repair function, which takes as input the invalid particle and returns as output the repaired one. At last the repaired output is returned. It explores the search space of possible solutions and returns the 3D structure with low energy. It starts with a set of randomly created potential solutions or particles gathered in a swarm. These particles are evaluated against their energy function values. At each iteration this swarm is updated using a function which updates the velocity of the particle. This operator's work is to discover new places of the search space to get the optimal solutions.

PSO for Identifying Gene Regulatory Networks

PSO can be utilized very effectively to solve the GRN identification problem. Each particle may represent the real valued expression levels of all the genes. Each gene has a specific expression level for another gene; thus a total of N genes correspond to N2 expression levels. Fitness of the particles may be computed from the absolute error with generated expression pattern (sum of all expressions) from the target expression pattern.

Role of PSO in Some Other Problems Relevant to Computational Biology

PSO for Feature Selection Using Gene Expression Data

Gene expression data represents the state of a cell at a molecular level, play an important role in medical diagnosis. One of the characteristics of gene expression data there are higher number of genes but comparatively smaller sample size. For which many researchers are presently learning how to select genes effectively before using a classification method to decrease the predictive error rate. In general, gene selection is based on two aspects: one is to find a group of genes having same functions and a close relationship, the other finds the smallest set of genes that can provide meaningful diagnostic information for disease prediction without compromising accuracy. Feature selection uses relatively fewer features since only selective features need to be used which in turn improve predictive error rate. The available training data are less in cancer classification, which is a challenge for some classification methodologies. So a PSO can be used here for doing the work of feature selection from a large number of features. An Improved binary particle swarm optimization (IBPSO) can be used in this case to do feature selection (4. In a binary PSO the individual particles can be represented as binary (0/1) forms which shows the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

straightforward "yes/no" choice of whether a feature needs to be selected or not. The changes in particle velocity can be interpreted as a change in the probability of finding the particle in one state or another. Since this change is a probability which is limited to a range of {0.0–1.0} (Chuang, Chang, Tu & Yang, 2008).

PSO for Rule Extraction in Medical Diagnosis

While applying conventional neural networks to the medical field it is very difficult to interpret appropriately, in a physically meaningful way, because the learned knowledge is numerically encoded in the trained synaptic weights. So some authors tried to generate if- then rules for transforming the synaptic weights in a neural network into a set of crisp rules. But after training, it is found that some if-then rules are present which are ineffective for classifying the positive instances. In this case a PSO can be used to reduce the rules which are mostly effective for classifying the positive data samples generated by a trained Neural Network, while the recognition performance will not be degraded, rather be improved. The membership function here chosen which measures the degree with which a data pattern, x, belongs to an input space. Initially the crisp rules generated are used as the particles, than the fitness function is evaluated to choose the best out of them (Hsieh, Chun & Wang, 2014).

PSO for Heating System Planning (HSP)

The whole world today faces serious energy and environment problems, with the increase of greenhouse gas emissions in the atmosphere the environments have already reached concerning levels in terms of their potential to cause climate change. Air pollution, acid precipitation, and stratospheric ozone depletion are other serious environmental concerns. An important action to address energy and environmental challenges lies in the intelligent and efficient use of energy, including reducing energy waste and using low- carbon fuels. Sustainable development of heating system requires application of planning procedures, which includes optimization of both demand and supply sides of heating. Here PSO algorithm can be used to design models for owners in making sustainable heating system investment decisions and to improve their decision-bases for administration. The objective function in HSP is the total heat production coast, which has to be minimized, for which PSO can be applied (Ma, Yu, & Hu, 2013).

Cancer Classification by BPNN-PSO

In medical field cancer is one of the major researches. Accurate prediction of various tumour types has huge value in offering better treatment possibilities and toxicity minimization on the patients. DNA microarrays performs biologist to calculate the expression of thousands of genes continuously on a small chip. These microarrays perform the large amount of data and proposed methods are needed to analyze them. The large amount of data needs more time for training when they are passed through a classifier. So to speed up the training time PSO is used. Searching process is started from initializing a group of random inputs. First, all the particles are updated, until a new generation set of particles are generated, and then those new particles are used to search the global best position in the solution space (Vimaladevi & Kalaavathi, 2014; Cho & Ryu, 2002).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTION

One of the foremost important problems associated with all population-based search algorithms is that they provide satisfactory results but there is no heuristic algorithm that could provide a superior performance than others in solving all kind of optimizing problems. In other words, an algorithm may solve some problems better and some problems worse than others (Rashedi, Nezamabadi-pour & Saryazdi, 2009). Hence, proposing new high performance heuristic algorithms are helpful. Recently some emerged techniques like gravitational search algorithms (GSA) can be used in place of PSO. This technique is different from PSO in many ways. In both GSA and PSO the optimization is obtained by agents' movement in the search space, but the strategy of movement is different. Some important differences are:

1. In PSO the direction of an agent is calculated using only two best positions, $pbest_i$ and gbest. But in GSA, the agent direction is calculated based on the overall force obtained by all other agents.
2. PSO uses a kind of memory for updating the velocity (due to $pbest_i$ and gbest). However, GSA is memory-less and only the current position of the agents plays a role in the updating procedure.
3. These advantages of GSA on PSO can be best experimented with the biological problems discussed in this chapter. Recently, an explosive growth is seen in the genomic sequences, proteins in available form, so, managing and manipulating these data and deriving the information accurately in a stipulated time bound has emerged as a challenging and fascinating field of science. So always there is a need to find suitable method and technology to handle big amount of data and the algorithms to process this information in an effective way. This may possibly be handled with suitable parallelization of algorithms. Hybridizing the intelligent soft computing tools with the heuristic swarm intelligence techniques in a parallel environment can be thought of as the future directions to find better solutions to address the above cited problems.

CONCLUSION

SI algorithms have recently gained wide popularity among the researchers, for their remarkable ability in obtaining nearly optimal solutions to a number of NP hard, real world search problems. A survey of the bioinformatics literature discloses the fact that the field has a surplus of problems that needs fast and vigorous search mechanisms. Problems coming under this group include (but are not limited to) the multiple sequence alignment (MSA), protein secondary and tertiary structure prediction, protein ligand docking, promoter identification and the reconstruction of evolutionary trees. Classical deterministic search algorithms and the derivative based optimization techniques are of no use for them as the search space may be massively large and discontinuous at certain parts. SI presents a collection of multi-agent parallel search techniques that can be effectively used for solving bioinformatics related problems of this sort. In this chapter we have tried to address some of the bioinformatics problems and their solutions using the computational intelligence. Then one of the popular and simple swarm intelligence based technique, the PSO has been discussed in various computational biology problems. But this field is not limited to the above. There is always a scope to innovate and renovate a better algorithm to produce a superior performance which will be suitable for answering the problems belonging to this domain.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Aziz, N. A. A., Mubin, M., Mohamad, M. S., & Aziz, K. A. (2014). Scientific World Journal, 2014.

Babu, M. M. (2004). *An Introduction to Microarray Data Analysis.* Retrieved from www.mrc-lmb.cam.ac.uk/genomes/madanm/microarray

Bai, Q. (2010). Analysis of Particle Swarm Optimization Algorithm. *Computer and information science, 3.*

Belal, M., Gaber, J., El-Sayed, H., & Almojel, A. (2006). Swarm Intelligence. In Chapman & Hall (Eds.), Handbook of Bioinspired Algorithms and Applications (Vol. 7).

Bharathi & Natarajan. (2010). Cancer Classification of Bioinformatics data using ANOVA. *International Journal of Computer Theory and Engineering*, *2*(3), 369–373.

Blum, C., & Merkle, D. (2008). *Swarm Intelligence – Introduction and Applications. Natural Computing*. Berlin: Springer.

Casadio, R., Compiani, M., Fariselli, P. & Vivarelli, F. (1995). Predicting free energy contributions to the conformational stability of folded proteins. *Journal of Intelligent System Molecular Biology*, 3, 81–88.

Chan, Z. S. H., Havukkala, I., Jain, V., Hu, Y., & Kasabov, N. (2008). Soft computing methods to predict gene regulatory networks: An integrative approach on time-series gene expression data. *Journa*l *of Applied Soft Computing*, *8*(3), 1189–1199. doi:10.1016/j.asoc.2007.02.023

Cheng, J. & Baldi, P. (2007) Improved residue contact prediction using support vector machines and a large feature set. *BMC Bioinform*, 8, 113.

Cho, S. B., & Ryu, J. (2002). Classifying gene expression data of cancer using classifier ensemble with mutually exclusive features *Proceedings of the IEEE*, *90*(11), 1744–1753. doi:10.1109/JPROC.2002.804682

Chuang, L. Y., Chang, H. W., Tu, C. J., & Yang, C. H. (2008). Improved binary PSO for feature selection using gene expression data. *Computational Biology and Chemistry*, *32*(1), 29–38. doi:10.1016/j.compbiolchem.2007.09.005 PMID:18023261

Chuang, L. Y., Chang, H. W., Tu, C. J., & Yang, C. H. (2008). Improved binary PSO for feature selection using gene expression data. *Computational Biology and Chemistry*, *3*(1), 29–38. doi:10.1016/j.compbiolchem.2007.09.005 PMID:18023261

Clote, P., & Backofen, R. (2000). *Computational Molecular Biology*. John Wiley & Sons Ltd.

Couzin, I. D., Krause, J., James, R., Ruxton, G. D., & Franks, N. R. (2002). Collective Memory and Spatial Sorting in Animal Groups. *Journal of Theoretical Biology*, *218*(1), 1–11. doi:10.1006/jtbi.2002.3065 PMID:12297066

Cui, Y., Chen, R. S., & Hung, W. (1998). Protein folding simulation with genetic algorithm and supersecondary structure constraints. *Proteins*, *31*(3), 247–257. doi:10.1002/(SICI)1097-0134(19980515)31:3<247::AID-PROT2>3.0.CO;2-G PMID:9593196

Das, S., Abraham, A., & Konar, A. (2008). Swarm Intelligence Algorithms in Bioinformatics, *Springer* [SCI]. *Studies in Computational Intelligence*, *94*, 113–147. doi:10.1007/978-3-540-76803-6_4

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Das, S., Panigrahi, B. K., & Pattnaik, S. S. (2009). Nature-Inspired Algorithms for Multi-objective Optimization. In Handbook of Research on Machine Learning Applications and Trends: Algorithms Methods and Techniques (Vol. 1, pp. 95–108). Hershey, PA, USA: IGI Global.

Dash, N., Priyadarshini, R., & Misra, R. (2015). An Artificial Neural Network Model to Classify Multinomial Datasets with Optimized Target Using Particle Swarm Optimization Technique, *Springer-Smart Innovation*. *Systems and Technologies*, *31*, 355–364.

Dembele, D., & Kastner, P. (2003). Fuzzy c-means method for clustering microarray data. *Bioinformatics (Oxford, England)*, *19*(8), 973–980. doi:10.1093/bioinformatics/btg119 PMID:12761060

Dorigo, M. (1992). Optimization, learning and natural algorithms (in Italian) [Ph.D. Thesis]. Dipartimento diElettronica, Politecnico di Milano, Italy.

Dorigo, M., Bonabeau, E., & Theraulaz, G. (2000). Ant algorithms and stigmergy. *Future Generation Computer Systems*, *16*(8), 851–871. doi:10.1016/S0167-739X(00)00042-X

Dorigo, M., Maniezzo, V., & Colorni, A. (1991). Positive feedback as a search strategy (Tech. Report 91-016). Dipartimento di Elettronica, Politecnico di Milano, Italy.

Engelbrecht, A. P. (2002). *Computational Intelligence: An Introduction*. England: John Wiley & Sons.

Goel, N., Singh, S., & Aseri, T. C. (2013).A Review of Soft Computing Techniques for Gene Prediction. Hindawi Publishing Corporation.

Haykin, S. (1999). *Neural networks—a comprehensive foundation*. Englewood Cliffs: Prentice Hall.

Hsieh, Y.-Z., Su, M.-C., & Wang, P.-C. (2014). A PSO-based rule extractor for medical diagnosis. *Journal of Biomedical Informatics*, *49*, 53–60. doi:10.1016/j.jbi.2014.05.001 PMID:24835617

Karaboga & Basturk, B. (2007). A Powerful and Efficient Algorithm For Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm. *Journal of Global Optimization, Springer Netherlands*, *39*(3), 459471.

Karaboga, (2005) An Idea Based On Honey Bee Swarm for Numerical Optimization (Technical Report-TR06). Erciyes University.

Kennedy, J. & Eberhart, R. (1995). An Introduction Particle Swarm Optimization. *IEEE Transaction*.

Li, B., Zheng, C. H., Huang, D. S., Zhang, L., & Han, K. (2010). Gene expression data classification using locally linear discriminant embedding. *Computers in Biology and Medicine*, *40*(10), 802–810. doi:10.1016/j.compbiomed.2010.08.003 PMID:20864095

Lim, C. P., Jain, L. C., & Dehuri, S. (2009). *Innovations in Swarm Intelligence: Studies in Computational Intelligence* (Vol. 248). Springer. doi:10.1007/978-3-642-04225-6_1

Ma, R.J., Yu, N.Y., & Hu, J.Y. (2013). Application of Particle swarm Optimization Algorithm in the Heating System Planning Problem. *The Scientific World Journal.*

Ma, R. J., Yu, N. Y., & Hu, J. Y. (2013). Application of Particle Swarm Optimization Algorithm in the Heating System Planning Problem. The Scientific World Journal.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mansour, N., Kanj, F., & Khachfe, H. (2012). *Particle Swarm Optimization Approach for Protein Structure Prediction in the 3D HP Model*. Interdisciplinary Science Compute Life Science.

Marquez-Chamorroa, A. E. (2015). Soft computing methods for the prediction of protein tertiary structures: A survey. *Journal of Applied Soft Computing*, *35*, 398–410. doi:10.1016/j.asoc.2015.06.024

Mitra, S., & Hayashi, Y. (2006). Bioinformatics with Soft Computing. *IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews*, *36*(5), 616–635. doi:10.1109/TSMCC.2006.879384

Panigrahi, K., Shi, Y., & Lim, M.-H. (2011). *Handbook of Swarm Intelligence. Series: Adaptation, Learning, and Optimization* (Vol. 7). Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-17390-5

Perez-Rodriguez, J., & Garcia-Pedrajas, N. (2011). An evolutionary algorithm for gene structure prediction. *Journal of Industrial Engineering and Other Applications.*

Priyadarshini, R., Dash, N. & Rout, S. (2012). A Novel Approach for Protein Structure Prediction using Back Propagation Neural Network. *International Journal of Computer Science & Technology*, 3(2).

Rashedi, Nezamabadi-pour & Saryazdi. (2009). GSA: a Grviational Search Algorithm. *Information Sciences*, 6, 2232-2248.

Sarhan, A. M. (2009). Cancer Classification Based on Micro array Gene Expression Data Using DCT and ANN. *Journal of Theoretical and Applied Information Technology*, *6*(2), 208–216.

Schulze-Kremer, S. (2000). *Genetic algorithms and protein folding, Protein Struct* (pp. 75–222). Predict.

Vaishali, P.K. & Vinayababu, A. (2011). Application of Microarray Technology and Soft-computing in Cancer Biology: A Review. *International Journal of Biometrics and Bioinformatics*, 5(4).

Vimaladevi, M. & Kalaavathi, B. (2014). Cancer Classification using Hybrid Fast Particle Swarm Optimization with Back-propagation Neural Network, *International Journal of computer and communication technology*, 3(11).

Xiao, X., Dow, E. R., Eberhart, R. C., Miled, Z. B., & Oppelt, R. J. (2003). Gene Clustering Using Self-Organizing Maps and Particle Swarm Optimization. *Proc. of the 17th International Symposium on Parallel and Distributed Processing (PDPS'03)*, Washington DC. 10.1109/IPDPS.2003.1213290

KEY TERMS AND DEFINITIONS

Agent: An agent is something that acts in an environment.

Bioinformatics: Bio-informatics can be described as the application of computational methods to make biological discoveries and explorations.

Bio-Inspired Computing: It is field which helps to tackle complex real problems using computational methods based on biology.

Collective Behaviour: Behaviour of a group of social insects.

Fuzzy System: It is a component machine learning techniques which takes membership values within 0 to 1 unlike crisp sets.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gene: It is a portion of DNA molecule which serves as the basic unit of heredity.

Heuristic: Rule or method based on trial and method that helps us to solve problems quicker.

Neural Network: Neural network is a soft computing paradigm inspired by the working of human brain to inculcate intelligence in machines.

Optimization: A mathematical method to find the solution of a problem towards achieving better performance either in form of minimum or maximum under one or more given constraint.

Protein: Organic macro molecules composed of one or more amino acids.

Swarm Intelligence (SI): SI is one of the techniques which is inspired by nature and is a population based algorithm motivated by the collective behaviour of a group of social insects.

This research was previously published in the Handbook of Research on Computational Intelligence Applications in Bioinformatics edited by Sujata Dash and Bidyadhar Subudhi; pages 119-143; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

PSO for Data Classification: A Case Study

Problem Definition:

Here we are considering a simple classification problem or we can say a prediction problem, where we are given with an input dataset sample having some instances and their predefined class. The main task in this problem is to partition the group of input objects to their corresponding classes.

In this work the bench mark datasets taken from UCI learning repository for verification and validation of the PSO based model is chosen (Marquez-Chamorroa, Asencio-Cortes, Santiesteban Toca & Aguilar-Ruiz, 2015) A brief of the datasets are as follows

Dataset Description:

The dataset consists of 150 samples which consists of a set of Iris flowers, where the goal is to predict three classes based on sepal (green covering) length and width, and petal (the flower part) length and width. This is the most popular and the simplest of the classification data containing three predictable attributes such as setosa, versicolor and virginica.

```
Input: iris dataset
Output: The instance and its corresponding class
Fitness function chosen is the minimum distance among the features.
Matlab Code:
 clc;
clear all;
irisdata;        % loading iris data into the work space
x=iris(:,1:4); % assigning instances to variable x
v=rand(1,4)-.5;
n1=1;n2=1;n3=1;
r1=.4;             %assigning constant value
r2=.5;
c2 =1.4;          % PSO parameter C1
c1 = 1.4;         % PSO parameter C2
w =0.73;         % PSO momentum or inertia
cd1=x(5,:);     %Randomly assigning data to one calss
cd2=x(65,:);
cd3=x(105,:);
n=1
while n<=5
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
for i=1:40
    t1=x(i,:);
    for j=1:40
        d1(j)=sqrt((sum(x(i,:)-cd1)).^2);
    end
    [m ind]=min(d1);
    g=x(ind,:);
        v = w *v + c1*r1*(t1-cd1) + c2*r2*(g-cd1);                    % ve-
locity updation
        cd1=cd1+v;
end
n=n+1;
cd1
end
n=1;
while n<=5
for i=51:90
    t1=x(i,:);
    for j=51:90
        d2(j)=sqrt((sum(x(i,:)-cd2)).^2);
    end
    [m ind]=min(d2);
    g=x(ind,:);
        v = w *v + c1*r1*(t1-cd2) + c2*r2*(g-cd2);
        cd2=cd2+v;
end
n=n+1;
cd2
end
n=1;
while n<=5
for i=101:140
    t1=x(i,:);
    for j=101:140
        d3(j)=sqrt((sum(x(i,:)-cd3)).^2);
    end
    [m ind]=min(d3);
    g=x(ind,:);
        v = w *v + c1*r1*(t1-cd3) + c2*r2*(g-cd3);
        cd3=cd3+v;
end
n=n+1;
cd3
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Simulation Accuracy of Iris Dataset

Number of classes	Number of Simulation Result Accurate up to				Total no of test samples
	4 decimal places	3 decimal places	2 decimal places	1 decimal places	
Class 1	10	10	10	10	10
Class 2	10	10	10	10	10
Class 3	09	09	09	09	10
Percentage of Accuracy	98%	98%	98%	98%	Total=30

```
end
 cd1
cd2
 cd3
```

Note: At first the data set will be divided into two sets. The first set will contain 2/3rd of the data which will be treated as training data and rest of the 1/3rd of the data will be treated as the test data. The output of this program will give the value of cd1, cd2 and cd3 which belong to class-1, 2 and 3 respectively for the training dataset. For the testing purpose any classifier can be used .Suppose an artificial neural network will be used as a classifier, the network will be trained by the calculated optimized features using PSO and the test data can be simulated by the classifier.

The Table 1 given below shows the simulated output of iris dataset. The results are compared by accuracy percentage and accuracy up to number of decimal places. Here 120 samples are taken as training and rest 30 are taken as test set.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 29
Semantic Reconciliation of Electronic Health Records Using Semantic Web Technologies

Karim El Guemhioui
University of Quebec in Outaouais, Canada

Steven A. Demurjian
University of Connecticut, USA

ABSTRACT

In this paper, the authors present an approach to reconcile the semantics of distinct medical terms found in personal health records (PHRs - that store data controlled by patients) and electronic medical records (EMRs - that store data controlled by providers) that are utilized to describe the same concept in different systems. The authors present a solution for semantic reconciliation based on RDF and related semantic web technologies. As part of the solution, the authors utilize a centralized repository of ontologies to: uniformly interrogate the medical coding systems in which those terms are defined, extract all of their published synonyms, and save the results as RDF triples. The final step in the process is to employ a reasoner to infer non-explicit synonymy among those terms, hence evidencing the underlying semantics to the PHR and EMR systems for possible further processing.

1. INTRODUCTION

In this age where smartphones are ubiquitous, and computer apps geared towards personal health issues proliferate, people are becoming more and more self-conscious of their health and incline to resort to these apps to keep track of their medical condition.[1] As a result, many of them end up assembling in apps health and fitness data from different sources that include several personal health records (PHRs)[2] that maintain data controlled by patients and electronic medical records (EMRs) that maintain data managed by medical providers. Patients utilize the mobile health (mHealth)[3] apps that seem to deal with distinct and unique health issues while in fact they are referring to the same health condition. While a medical

DOI: 10.4018/978-1-7998-1204-3.ch029

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

provider is easily able to identify related different terms in PHRs and/or EMRs referring to the same condition (i.e., problem), this exercise is less obvious for most individuals (whose medical knowledge is usually superficial), and outright impossible for the machine hosting these electronic records and the computer applications we rely on to make some sense out of this data.

The multiplication and affordability of electronic mobile devices (smartphones, phablets, tablets, "Fitbits", wearable healthcare devices, etc.), and their embracing of the technologies by an increasing number of diverse stakeholders (e.g., patients, family members, home health care aids, therapists, etc.) are paving the way for a shift in the gathering and integration of personal healthcare data from PHRs and EMRs coupled with its readiness for the various healthcare and medical service providers. Therefore, the integration of PHRs and EMRs will play a pivotal role in this move to a safer, centralized, more efficient, consumer-driven healthcare system. Just to give a simple example: mHealth applications that provide access to health and fitness data, gathered from multiple PHRs and EMRs, will allow us to have our medical history available in case of emergency, when traveling, seeking a second opinion, or switching health insurance companies (HealthIt, 2014).

To accomplish this goal, we are starting to see official efforts focused on encouraging provider organizations to give consumers access to their own data; e.g., the Blue button initiative (HealthIT, 2016). More and more sophisticated PHR apps are also starting to support access to and/or import of electronic medical records (EMRs) or electronic health records (EHRs)[4]. A new term has even been coined: a tethered (or connected) PHR to designate a PHR linked to a specific healthcare organization's electronic health record (EHR) system or to a health plan's information system (Daglish & Archer, 2009). With a tethered PHR, patients can access their own records through a secure portal and see, for example, the trend of their lab results over the last year, their immunization history, or due dates for screenings.

While there is no agreed upon standard for a PHR, a significant amount of standardization efforts has been done regarding the annotation, classification, and tagging of health and fitness data in EMRs and EHRs. Among the most widespread standards, at least in North America, are: HL7 v2.x (1989), CDA (2005), LOINC[5], SNOMED[6], UMLS[7], and FHIR (2014). PHRs accessing/downloading these remote electronic data are therefore exposed to their underlying format. This paper focuses on the data in the PHR obtained/constructed from remote EMRs/EHRs sources, and leverages existing standardization efforts to propose an approach to the semantic reconciliation of some critical information appearing in PHRs (and especially tethered PHRs). As for the medical terms used in these EHRs to describe a health condition, a treatment, a medication, etc., we are assuming that they comply with agreed upon terminologies defined in specialized thesauri, controlled vocabularies, reference books, dedicated ontologies, etc., which we will call generically "medical coding systems" in the remainder of this paper. Also, for the sake of clarity, we are limiting ourselves in this work to clinical terms describing a medical condition (i.e., an observed problem).

This paper presents a solution for semantic reconciliation that utilizes a centralized repository of ontologies to identify the relevant terms, extract all published synonyms, and save the results as RDF triples. To facilitate the discussion, we present a scenario that illustrates a rather common issue in PHRs, namely, that the same medical condition is being designated by different medical terms in different EMRs/EHRs, creating therefore a semantic hiatus and preventing appropriate machine processing. This scenario is utilized to demonstrate our approach to semantically reconcile these related terms, whether they have been chosen from the same or from distinct medical coding systems. The proposed solution is built around the Resource Description Framework (RDF) and its related standards RDFS, OWL, SPARQL, etc. (Allemang and Hendler, 2011). Note that the presented approach straightforwardly extends to terms dealing with others issues than medical condition (e.g., medication, treatment, lab results).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The remainder of this paper is organized into five sections. In Section 2, we provide some background information and definitions relevant to our work including: the different types of health information technology systems (PHRs, EMRs, EHRs) and the role of health information exchange among these systems. In Section 3, we explain the challenge resulting from the multiplication of healthcare concept definitions of medical nomenclature, and provide a use case scenario that will serve to illustrate our approach. Section 4, the main focus of this paper, details and explains our approach to reconcile semantic data in EMRs/EHRs in order to reconcile terms across multiple systems to present a unified version to patients via their mHealth apps. Section 5 discusses the semantic harmonization issue that involves the ability to achieve the desired reconciliation utilizing the new FHIR (2014) standard. Section 6 concludes this paper.

2. BACKGROUND AND RELATED WORK

A personal health record (PHR) is a collection of information about an individual's health that is maintained by the individual. A PHR system is an electronic application used by patients to maintain and manage their health information in a private, secure, and confidential environment. Thus, the PHR concept is patient centric in the sense that its management remains primarily the responsibility of the patient. Genitsaridi, Kondylakis, and Koumakis (2015) have identified a number of five functional criteria that an ideal PHR system should meet:

1. **Problem, Diagnosis and Treatment (PDT) Basic:** This category is of most importance to our work and will be detailed next.
2. **Self-Health Monitoring:** This category includes functions that help the patient monitor her own health status. The data recorded by functions in this category are unofficial and generated by the patient.
3. **Communication Management:** This category includes functions that provide services such as appointment scheduling, appointment reminders, recording of healthcare professionals' information (name, phone, address ...), etc.
4. **Security and Access Control:** This category involves security (authorization and authentication) as well as access control (privileges and permissions).
5. **Intelligence Factors:** This category includes all functions that illustrate intelligent behavior, such as intelligent data presentation, intelligent data export, intelligent system alerts, interaction with other electronic health record (EHR) systems, etc.

   The first category includes functions that provide the following services:
   - Recording of a patient's problem or an incident.
   - Recording of the diagnosis process related data such as diagnostic tests (e.g., blood, laboratory tests, imaging tests, EKGs, etc.), diagnostic procedures (e.g., therapy, surgery, etc.), diagnostic material (e.g., allergies, family history, immunizations, etc.), and diagnostic results.
   - Recording of the treatment process related data such as treatment procedures, medications and hospitalizations.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The data that are recorded by this category's functions are official facts generated by medical providers in the form of Electronic Medical Records (EMRs) or Electronic Health Records (EHRs). In this paper, we are mainly interested in PHR systems that support this category's functions. A PHR differs from an EMR or an EHR, which are owned and stored by medical providers. An EMR is a digital version of a paper chart that contains the standard medical and clinical data gathered in one provider's office. Electronic health records (EHRs) go beyond the data collected in the provider's office and include a more comprehensive patient history. For example, EHRs are designed to contain and share information from all providers involved in a patient's care. EHR data can be created, managed, and consulted by authorized providers and staff from across more than one health care organization (HealthIT, 2014). Unlike EMRs, EHRs also allow a patient's health record to move with them, to other health care providers, specialists, hospitals, nursing homes, etc.

Anytime a PHR holder moves, switches medical providers, or seeks medical treatment during a vacation, he/she will be asked to provide his/her health history which, usually, is difficult to remember. The availability of a PHR allows an individual to more easily answer any questions such as vaccinations and medications/dosages. A PHR can also be a live saver in an emergency situation. For example, if a person were to experience a health emergency at work or on a trip with friends, his/her companions would hardly be able to answer questions about his/her health history and current medical providers. Furthermore, in such a time of crisis, it can be difficult to remember important and potentially life-saving information, whereas a PHR can provide the health information promptly and reliably.

In order for the electronic health information exchange (HIE) to succeed, it is imperative that the medical providers agree on an exchange standard. Several exchange standards (e.g., HL7 v2.x (1989), CDA (2005), LOINC, SNOMED, UMLS; and FHIR (2014)) have been developed over the years, each new standard purporting to overcome the shortcomings of the previous one. The latest one is FHIR – Fast Healthcare Interoperable Resources (Bender & Sartipi, 2013), which builds on previous data format standards from the Health Level Seven International Organization, and is supposedly easier to implement. One of the goals of FHIR is to facilitate interoperation between legacy healthcare systems, to make it easy to provide healthcare information to medical providers and individuals on a wide variety of devices from computers to tablets to smartphones, and to allow third-party application developers to provide medical applications which can be easily integrated into existing systems. A more detailed review of these electronic exchange standards is beyond the scope of our work, and has been abundantly addressed in the scientific literature (Eichelberg et al., 2005; Smits et al., 2015).

3. MEDICAL CODING SYSTEMS AND UNDERLYNG SEMANTICS

Healthcare data electronically exchanged among medical providers can take up varying formats and use non-compatible terminology. Ultimately, some of this data ends up in PHRs that appear to describe distinct medical problems, while in fact they convey the same semantics. This regrettable situation can be retraced to the existence of a multitude of standardized terminologies from which to choose those medical terms. Section 3.1 briefly reviews the nomenclature in the medical field, while Section 3.2 introduces a use case scenario of healthcare to be utilized for examples throughout the rest of the paper.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.1. The Medical Terminology Bazaar

For health information exchange (HIE), as discussed in Section 2, to succeed agreeing on an exchange standard is necessary but not sufficient. There must also be some agreement on the clinical terminology used in those electronic messages. Unfortunately, there exist several competing medical coding systems with distinct or overlapping terms to designate the same concept or a closely related one; and medical providers familiarized with one coding system are very reluctant to switch to another one. In this work, we limit ourselves to three ontologies (SNOMED-CT, CD-10, and OBO DOID) to illustrate our point regarding the profusion of terms of equivalent meaning within and among coding systems, and hence the semantic gap that arises.

The Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT, 2016) is a standardized, multilingual vocabulary of clinical terminology used by physicians and other health care providers for the electronic exchange of clinical health information. The standard CD-10-CM (2016) is the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD), a medical classification list by the World Health Organization (WHO). Lastly, the standard Human Disease Ontology (OBO DOID, 2016) is an ontology defined by the Open Biomedical Ontologies Foundry for describing the classification of human diseases organized by etiology. In most of these medical coding systems, there are usually several synonyms that are utilized to represent the same concept. To disambiguate the situation, a unique identifier is assigned to the concept and shared by all the synonyms. Furthermore, at least one of the terms is designated as the preferred one, to spur its use by the medical providers (the others are alternate terms). Finally, note that medical providers must comply with these medical coding systems in order for their EHRs to be intelligible and in the US to satisfy certification[8] for Health IT systems. Further, there is also the Center of Medicare and Medicaid Services Meaningful Use Stage 3 guidelines[9] that require all health information in EMRs and PHRs to meet certain requirements regarding standards and availability of data via cloud services. This means that there are efforts in the US to insure a commonality in standards.

3.2. A Use Case Scenario

The following scenario will be used to illustrate our approach in reconciling the semantics of critical data in EHRs. Let's consider a person, say Mr. John Smith, who likes to keep track of his medical history. For that purpose, he uses an application ("apps") that gathers all of his medical records on his smartphone.

1. June 15th, at around 3:45 pm, Mr. John Smith who is at home feels a severe chest pain. Because his family doctor is at least 40 min drive, and on the insistence of his wife, an ambulance is summoned that takes Mr. John Smith to the emergency room (ER). There, he is diagnosed by an intern with a "Cardiac infarction," and an EHR is produced.
2. The following day, Mr. John Smith consults his family doctor. A second EHR mentioning that Mr. John Smith has suffered from an "Attack - heart" is produced. Mr. John Smith is referred to a cardiologist for further treatment.
3. Eight days later, Mr. John Smith manages to be seen by the heart specialist who diagnoses a "Myocardial infarction" and produces a third EHR. The electronic message in FHIR format is shown in the next section (Figure 1).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. Ten days later, a follow up with the cardiologist results in a fourth EHR indicating that Mr. John Smith suffers from a "Microinfarct of heart."

Ultimately, all of these data from these EHRs end up in Mr. John Smith's PHR and are accessible via his mHealth application. The mechanics of the way that he obtains access to his personal EMR/EHR via his PHR are beyond the scope of this paper.

Though a person in the healthcare domain may, more or less easily, recognize that the first three medical records refer to 3 occurrences of the same incident, i.e., a "Myocardial infarction" (preferred term in the SNOMED-CT coding system), it would be desirable for other systems and mHealth apps to be able to do the same, specifically, to detect that the three medical terms employed are semantically equivalent. This semantic reconciliation is essential for the proper and consistent functioning of mHealth apps, their correct processing of the data they feed upon, and the accuracy and/or reliability of the results they present us with. Furthermore, the mHealth apps user, in our specific case, Mr. John Smith, should also be able to clearly and quite immediately relate the three concepts, regardless of whatever other functionalities the apps provides. Remember that though Mr. John Smith may be quite knowledgeable with respect to his medical situation, his use of mHealth apps being a clear indication of his commitment, it's not obvious under the described scenario that he would adequately relate the three incidents. One reason could be that his deteriorating health may have affected his cognitive functions. The fourth medical condition, in our scenario, will allow us to explore a more complex situation, since the term used is not a synonym of the first three. We believe that this scenario is not contrived and illustrates a very common problem in the medical/healthcare world which is the proliferation of terms to designate the same concept or a very closely related one.

4. AN APPROACH TO ESTABLISH SEMANTIC EQUIVALENCE OF MEDICAL CONCEPTS

This section is the core of the paper to present an approach that allows the semantic equivalence of medical concepts to be achieved. The discussion is organized into three parts. First, Section 4.1 discusses the way that relevant health data can be captured and represented as RDF triples. Next, Section 4.2 presents the usage of the Simple Knowledge Organization System (SKOS)[10] for knowledge representation based on the resource definition framework (RDF)[11] in conjunction with the SPARQL[12] database query language for data stored in RDF format. Finally, Section 4.3 explores an approach to infer more complex semantic relationships, such as the "Microinfarct of heart" in the fourth EHR of the scenario in Section 3.2 that will be reconciled with the synonymous terms in the other three EHRs.

4.1. Capturing Relevant Healthcare Data into Triples

EHRs/EMRs can come in different shapes with varying APIs, in essence, the data they encompass are usually stored in either relational or XML repositories with the ability to export information into HL7 for potential sharing. For semantic reconciliation, we would want to transition data retrieved from an EHR/EMR/PHR via XML into the RDF format. This allows data that is retrieved from the four different EHR sources in the Scenario of Section 3.2 to be uniformly stored in RDF format as triples: subject – property/predicate - object. This very uniform and regular structure will significantly ease the parsing

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and processing of the medical information we are interested in. Otherwise we would have to resort to complex parsing code/scripts to isolate the medical terms whose semantics we want to explicit.

A great deal of work has been done regarding the automatic conversion of XML files and/or relational tables into RDF triples (Breitling, 2009; Bohring, & Auer, 2005; Sahoo et al., 2009). Though the exact mechanics of these transformations are beyond the scope of this paper, we can reasonably assume that the automatic conversion from the four EHRs of the scenario into RDF and its related standards, such as RDFS and OWL, is fairly achievable. Furthermore, if the data portion we are mostly interested in (i.e., imported EHR data) is acquired using a standardized messaging format for exchanging information electronically, there probably already exists syntactic transformations to convert it into RDF triples. While for some imported standard messages, such as HL7 v2.x, the transformations can be rather complex and/or found wanting (Prasser et al., 2012; Kawazoe, Imai, & Ohe, 2016); for some others, such as FHIR messages, substantial work is being done in defining mappings to RDF (and its related standards: RDFS, OWL). We believe that automated transformation and validation tools should begin to appear.

As an example, the FHIR Working Group of the HL7 organization has a section dedicated to RDF[13]. As an illustration, we provide below an example of a FHIR message in XML format (one of the two fully supported formats, the other being JSON) and its translation into a FHIR/RDF compliant format. The original FHIR/XML message (slightly truncated for brevity) in Figure 1 represents the EHR produced by the cardiologist following his first examination of Mr. John Smith (the third medical condition in our scenario). The FHIR message clearly identifies (highlighted by us in blue): the coding system used, the unique identifier for the medical condition reported, and a textual value.

The translation of this message into FHIR/RDF, using the Turtle syntax, is shown in Appendix A. The Turtle code provided has been validated using the Apache Any23 tool[14]. Note that this same tool (though there are many others) allows one to convert the terse RDF Triple Language notation (Turtle) into the simpler N-Triples notation which is even easier for software to parse.

4.2. SKOS and SPARQL to the Rescue

To support the semantic reconciliation process, we seek to transition from RDF triples to the usage of the Simple Knowledge Organization System (SKOS) for knowledge representation based on the resource definition framework (RDF) and the SPARQL database query language. As already mentioned, our objective is to make possible for the machine/apps to infer that the first 3 EHRs of our scenario in Section 3.2 are in fact referring to the same medical condition. A naive and tedious approach would be to take each medical condition mentioned in our scenario, look it up in the coding system in which it is defined and retrieve all of its synonyms, if any. The next step would be to compare each one of them with all the synonyms for the other medical conditions to see if there is a match. If yes, then the involved EHRs are referring to the same medical condition; if not, we cannot at this stage reject since there might still exist some close relatedness between the different data stored in the three EHRs, as will be illustrated in Section 4.3 with the 4th medical condition of our scenario in the fourth EHR. There is a need for us to leverage both SKOS and SPARQL in order to come up with a unified process to extract and compare data from different EHRs. Without these technologies, we would need to formulate a correct lookup query for each EHR utilizing the specific medical coding system of that EHR. That would be problematic for multiple EHRs with different coding systems that are not compatible. Therefore, our approach is intended to support the issuance of generic queries to arbitrary medical coding systems (thesauri, ontologies,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A FHIR/XML message for third medical condition in scenario

```
<Condition xmlns="http://hl7.org/fhir">
  <id value=" f001"/>
  <text>
    <status value="generated"/>
    <div xmlns="http://www.w3.org/1999/xhtml">Myocardial infarction,
June 15, 2016</div>
  </text>
  <patient>
    <reference value="Patient/f003"/>
    <display value= "Mr. John Smith"/>
  </patient>
  <code>
    <coding>
      <system value="http://snomed.info/sct"/>
      <code value="22298006"/>
      <display value=" Myocardial infarction "/>
    </coding>
    <text value=" Myocardial infarction "/>
  </code>
 <category>
   <coding>
     <system value="http://snomed.info/sct"/>
     <code value="439401001"/>
     <display value="diagnosis"/>
    </coding>
 </category>
   <clinicalStatus value="active"/>
   <verificationStatus value="confirmed"/>
 <severity>
    <coding>
      <system value="http://snomed.info/sct"/>
      <code value="6736007"/>
      <display value="Moderate"/>
      </coding>
    </severity>
  <onsetDateTime value="2016-06-15"/>
</Condition>
```

etc.) to fetch the medical concepts that we are interested in, without having to concern ourselves with the specific structure/format of each coding system, nor the knowledge of which coding system actually contains the information we are after. SKOS and SPARQL can facilitate this process.

Fortunately, this is possible (with very reasonable limits) due to the efforts of the National Center for Biomedical Ontology (NCBO) who has developed BioPortal[15]. BioPortal is an open and comprehensive repository of biomedical ontologies (more than 400) with services that provide access to those ontologies (SPARQL endpoint, REST API) and a set of tools for working with them (browser, text annotator, etc.) (Noy et al., 2009; Salvadores et al., 2012). The two aspects that are of most interest to us are: (1) the hosted ontologies (i.e., medical coding systems) that have been transformed into RDF triples from their original formats (OWL, OBO, UMLS/RRF, ...) and asserted into a triple store; and, (2) metadata that have been injected on top of original related terms by using the predicates of SKOS, skos:prefLabel and skos:altLabel[16] (Miles et al., 2005), to single out the preferred name and the synonyms, respectively (usually, as identified by those who submitted the ontology to the portal).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Let's consider each one of the medical conditions mentioned in our scenario and demonstrate the way that it is possible to establish their semantic equivalency or relatedness, if any. We start with the 3rd medical condition for which a FHIR message correctly filled out (see Figure 1 again) has been provided. Since the coding system and the unique identifier are known, the trivial SPARQL query, shown in Figure 2, provides us with all the valid names the reported medical condition is also known under this coding system. For clarity, we will characterize all semantically equivalent terms as synonyms (i.e., the fact that one of them is labeled as preferred term is inconsequential for our purpose).

It is important to note that to write this query, there is no need to be knowledgeable about SNOMED-CT, its internal organization, what's the name of the property for an equivalent term, etc. We are only relying on two well known properties of the widely published ontology SKOS namely, skos:prefLabel and skos:altLabel. We are also assuming that the mappings of these 2 properties to their counterparts in the BioPortal ontologies have been done correctly; and note that this issue is a different research problem out of the scope of our paper.

The intent of the process is to arrive at a list of synonymous terms, as shown in Figure 3, where the term used by the cardiologist turns out to be the preferred name (after all she is an expert), and it has 10 other alternate names, for a total of 11 synonyms for the same medical condition (and just for US English).

Figure 2. A SPARQL query

```
PREFIX skos:  <http://www.w3.org/2004/02/skos/core#>
PREFIX sno:   <http://purl.bioontology.org/ontology/SNOMEDCT/>
SELECT DISTINCT  *
WHERE {
          { sno:22298006  skos:prefLabel  ?prefLabel . }
          UNION
          { sno:22298006  skos:altLabel    ?synonym . }
 }
```

Figure 3. Synonymous terms of myocardial infarction

prefLabel	synonym
"Myocardial infarction"@EN	---
---	"Myocardial infarct"@EN
---	"MI - Myocardial infarction"@EN
---	"Heart attack, NOS"@EN
---	"Infarction of heart"@EN
---	"Myocardial infarction, NOS"@EN
---	"Infarction of heart, NOS"@EN
---	"Heart attack"@EN
---	"Cardiac infarction, NOS"@EN
---	"Myocardial infarction (disorder)"@EN
---	"Cardiac infarction"@EN

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

While the information in Figure 3 is very informative for a human, it is not directly "processable" by the machine. What we really need is to have the query results formatted as triples so that they can be stored locally for further processing. This is easily achievable in SPARQL by using the CONSTRUCT keyword as shown in Figure 4.

This query in Figure 4 creates a set of triples, in which the subject is always the unique identifier of the concept we are interested in, the predicate is some property that conveys the idea of synonymy, and the object is any term by which the concept is also known. Instead of creating our own property (e.g., prefixForOurNameSpace:isSemanticallyEquivalentTo) which wouldn't be recognized by the BioPortal, we adopt a property already known to the system, namely, the owl:sameAs[17]. Note that the controversy that has surrounded this predicate (Halpin et al., 2010) is insignificant in our context. The result is shown in Figure 5 and stored locally. Note that for brevity, we have shortened the subject namespace http://purl.bioontology.org/ontology/SNOMEDCT to SNOMEDCT, and the predicate namespace http://www.w3.org/2002/07/owl to owl).

Next, we consider the first medical condition of our scenario, for which, obviously, we don't know the unique identifier. However, we assume that we know the coding system utilized (namely, SNOMED-CT). Again, though the term utilized for the medical condition (i.e., "Cardiac infarction") can easily be spotted by a human in the previous results, the machine doesn't know that yet. The SPARQL query in Figure 6 produces human readable results (akin to Figure 3), while the change shown in Figure 8 to this SPARQL query generates the RDF triples we need to store.

Figure 4. SPARQL query by CONTRUCT

```
PREFIX skos:  <http://www.w3.org/2004/02/skos/core#>
PREFIX owl:   <http://www.w3.org/2002/07/owl#>
PREFIX sno:  <http://purl.bioontology.org/ontology/SNOMEDCT/>
CONSTRUCT { sno:22298006 owl:sameAs  ?prefLabel ,
                                     ?synonym . }
WHERE {
        { sno:22298006  skos:prefLabel   ?prefLabel . }
        UNION
        { sno:22298006  skos:altLabel    ?synonym . }
 }
```

Figure 5. Results of the SPARQL Query in Figure 4

```
<SNOMEDCT/22298006> <owl#sameAs> "Myocardial infarction"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Myocardial infarct"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "MI - Myocardial infarction"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Heart attack, NOS"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Infarction of heart"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Myocardial infarction, NOS"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Infarction of heart, NOS"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Heart attack"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Cardiac infarction, NOS"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Myocardial infarction (disorder)"@EN.
<SNOMEDCT/22298006> <owl#sameAs> "Cardiac infarction"@EN .
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Since we don't know if the medical term we are looking up is a preferred term or an alternate one, we need to try both cases. The result produced by the query in Figure 6 is as follows (Figure 7).

To produce the result formatted as triples, we just need to replace the line of code "SELECT DISTINCT *" in Figure 6 with the code shown in Figure 8.

This substitution yields the following triples (namespaces shortened as explained earlier) which we also store locally:

```
<SNOMEDCT/22298006> <owl#sameAs> "Myocardial infarction"@EN .
<SNOMEDCT/22298006> <owl#sameAs> "Cardiac infarction"@EN .
```

In our approach, we are taking advantage of the support of regular expressions in SPARQL to search for exact matches. Of course if the search came empty, we could broaden our search criteria in several ways (case insensitive, empty spaces, substrings, etc.). The example in Figure 9 illustrates one possibility.

Figure 6. SPARQL query for human readable results in SNOMED-CT

```
PREFIX owl:  <http://www.w3.org/2002/07/owl#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT  *
FROM <http://bioportal.bioontology.org/ontologies/SNOMEDCT>
FROM <http://bioportal.bioontology.org/ontologies/globals>
WHERE {
          {
            ?termURI   a       owl:Class  ;
             skos:prefLabel    ?prefLabel  ;
             skos:altLabel     ?synonym  .
             FILTER (regex(?prefLabel, '^Cardiac infarction$'))
          } UNION {
            ?termURI   a       owl:Class  ;
             skos:prefLabel    ?prefLabel  ;
             skos:altLabel     ?synonym  .
             FILTER (regex(?synonym, '^Cardiac infarction$'))
          }
}
```

Figure 7. Result of query

termURI	prefLabel	synonym
<http://purl.bioontology.org/ontology/SNOMEDCT/22298006>	"Myocardial infarction"@EN	"Cardiac infarction"@EN

Figure 8. Change to SPARQL query of Figure 6 for machine consumption

```
CONSTRUCT { ?termURI   owl:sameAs    ?prefLabel  ,
                                     ?synonym  .
          }
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Using regular expressions to broaden search

```
FILTER (  ( CONTAINS  (UCASE(str(?prefLabel)), "CARDIAC")   &&
            CONTAINS  (UCASE(str(?prefLabel)), "INFARCTION"))
           ||
           ( CONTAINS  (UCASE(str(?synonym)), "CARDIAC")   &&
            CONTAINS  (UCASE(str(?synonym)),"INFARCTION"))  )
```

To continue the discussion, we proceed to the 2nd medical condition of our scenario in Section 3.2 which was diagnosed by the family doctor. This time not only the unique identifier is missing, but also the medical coding system used. A human observer may be tempted to conclude that this condition is the same as the first one (since the expression "Heart attack" appears as one of the synonyms for the first condition of scenario). But again, we would like the machine to infer such a result, which cannot be done based only on the two involved strings: "Heart attack" and "Attack – heart". Furthermore, not knowing which medical coding system this latter term was taken from, we cannot be sure of the specific medical condition it describes, the vagueness of the expression being often used as a catch-all for various heart related pathologies. The SPARQL query in Figure 10 provides an answer to our concerns.

Since we have not specified which ontologies to search (via the FROM keyword), the default behavior will be to explore all of the ontologies in the repository. However, the query in Figure 10 could not run its course since the SPARQL BioPortal endpoint is an experimental one (in beta version), for demo purposes, that cannot handle such huge search space. To overcome this limitation, we included in the search space a number of ontologies as shown in Figure 11.

Figure 10. SPARQL query for second medical condition

```
PREFIX  owl:   <http://www.w3.org/2002/07/owl#>
PREFIX  skos:  <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT   *
### No ontology to search is specified
WHERE {
          {
            ?termURI   a        owl:Class  ;
             skos:prefLabel    ?prefLabel  ;
             skos:altLabel     ?synonym  .
             FILTER (regex(?prefLabel, '^Attack - heart$'))
          } UNION {
            ?termURI   a        owl:Class  ;
             skos:prefLabel    ?prefLabel  ;
             skos:altLabel     ?synonym  .
            FILTER (regex(?synonym, '^Attack - heart$'))
          }
}
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. SPARQL with added Ontologies

```
PREFIX owl:  <http://www.w3.org/2002/07/owl#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT  *
FROM <http://bioportal.bioontology.org/ontologies/ICD9CM>
FROM <http://bioportal.bioontology.org/ontologies/ICD10CM>
FROM <http://bioportal.bioontology.org/ontologies/HP>
FROM <http://bioportal.bioontology.org/ontologies/DOID>
FROM <http://bioportal.bioontology.org/ontologies/TMO>
#FROM <http://bioportal.bioontology.org/ontologies/SNOMEDCT>
FROM <http://bioportal.bioontology.org/ontologies/globals>
WHERE { # same code as the previous query }
```

In Figure 11, the SNOMED-CT ontology had to be commented out since it was the one triggering the time out before the end of the search. We nevertheless made sure that we weren't missing anything by running the query on the SNOMED-CT alone, and it came empty[18]. Below is the result of our query and also a clear indication of the medical coding system used, namely the OBO DOID, also known as the Human disease ontology (see Figure 12).

As before, by replacing the line starting with SELECT (in Figure 11) with the following code:

```
CONSTRUCT  {  ?termURI    owl:sameAs     ?prefLabel,
?synonym  .
  }
```

```
<http://purl.obolibrary.org/obo/DOID_5844>
  <http://www.w3.org/2002/07/owl#sameAs>
   "myocardial infarction"^^<http://www.w3.org/2001/XMLSchema#string> .
<http://purl.obolibrary.org/obo/DOID_5844>
  <http://www.w3.org/2002/07/owl#sameAs>
    "Attack - heart"^^<http://www.w3.org/2001/XMLSchema#string> .
```

The 15 generated triples stored in our local machine must then be fed to an inference engine to evidence the existing semantic equivalence. After saving the newly inferred triples, we can run the SPARQL code in Figure 13 to produce a cleaned out output. The predicate owl:sameAs being symmetric and transitive, we have to filter out undesirable results (every term is same as itself). To keep just the textual values of the medical terms, we also filter out the unique identifiers.

Figure 12. Result of query

termURI	prefLabel	synonym
<http://purl.obolibrary.org/obo/DOID_5844>	"myocardial infarction"	"Attack - heart"

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 13. SPARQL code to remove redundant results

```
CONSTRUCT { ?a  owl:sameAs  ?b . }
WHERE {
        ?a  owl:sameAs  ?b  .
        FILTER (  (?a != ?b)  && (!isURI(?a)) && (!isURI(?b)) )
}
```

For demonstration purposes, we have resorted to a "quick and dirty" shortcut to clearly visualize the results, by loading our "mini ontology"[19] into the open-source ontology editor and framework Protégé, and running its default reasoner (FACT++). The system has then inferred that the first three medical conditions of our scenario were semantically equivalent, as shown in Figure 14 where the inferred terms found semantically equivalent, have a yellow background. Thought we had 15 triples, only the 13 distinct terms are shown (2 triples were asserting duplicate information).

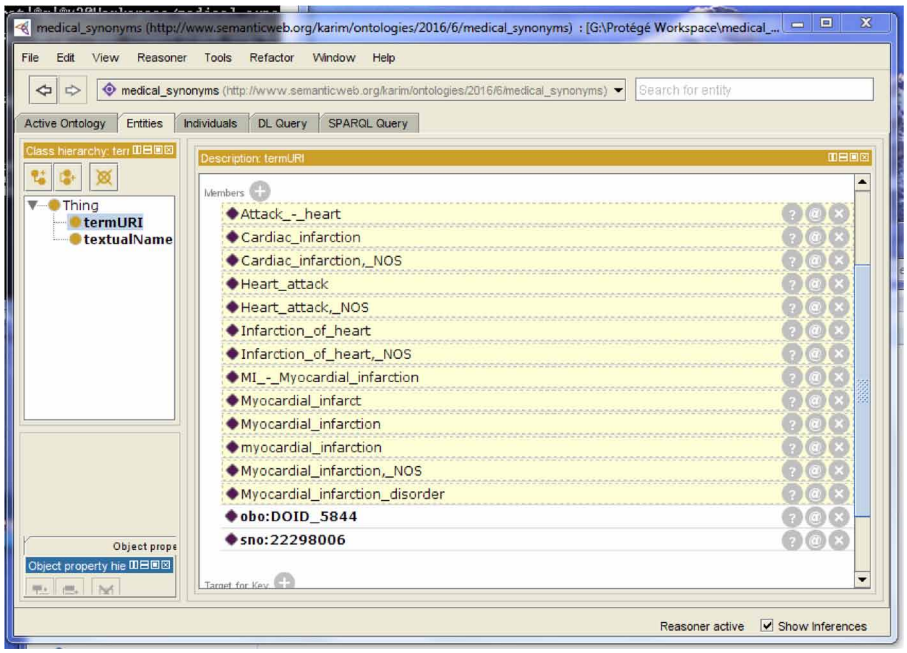
We can also, and more simply, verify semantic equivalence on a case by case basis.

```
ASK WHERE { "Cardiac infarction"   owl:sameAs  ?any . }
```

This query will yield TRUE because several matches will be found in our local triple store.

```
ASK WHERE { "Microinfarct of heart"   skos:sameAs  ?any . }
```

Figure 14. Visualizing the results in Protégé



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This query will yield FALSE because this medical condition never appears in our triple store as the subject or the object of the predicate owl:sameAs.

4.3. Inferring More Complex Semantic Relationships

This section infers more complex semantic relationships, such as the a "Microinfarct of heart" in the fourth EHR of the scenario in Section 3.2 that need to be reconciled with the synonymous terms in the other three EHRs. Specifically, we are interested in determining if the 4th medical condition of our scenario ("Microinfarct of heart") is closely related to the other three. Several kinds of explorations can be done with SPARQL. For brevity, we will restrict ourselves to an obvious one: is this concept a subclass of the concept depicted in the first three situations of our scenario (whose unique identifier is 22298006) as shown with the SPARQL query and result in Figure 15.

The result at the bottom of Figure 15 demonstrates that this last medical condition is a subclass of the previous one. Again, we have first tried to look for an exact match of the medical textual term used by the cardiologist. If the search had come empty, we would have relaxed the search criteria to ignore for example the letter case, etc. We notice also that this last medical condition has two synonyms. To generate the triple to store locally, we need to replace the FILTER line in Figure 15 with:

```
CONSTRUCT  { ?child  skos:broader   ?parent  . }

<http://purl.bioontology.org/ontology/SNOMEDCT/42531007>
       <http://www.w3.org/2004/02/skos/core#broader>
             <http://purl.bioontology.org/ontology/SNOMEDCT/22298006> .
```

For simplicity, we have related the child and parent concepts by using the **skos:broader** property, since the SKOS ontology has already been imported. We recall that this property means that the first concept (the child class) has a broader concept (the parent class).

Figure 15. SPARQL query and result for fourth medical condition

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX  sno: <http://purl.bioontology.org/ontology/SNOMEDCT/>
SELECT DISTINCT  ?childPrefLabel  ?synonym  ?parentPrefLabel
WHERE {
        GRAPH  <http://bioportal.bioontology.org/ontologies/SNOMEDCT>  {
              sno:22298006  skos:prefLabel     ?parentPrefLabel  .
              ?child            rdfs:subClassOf    ?parent  ;
                                kos:prefLabel      ?childPrefLabel ;
                                skos:altLabel      ?synonym .
        FILTER (regex(?childPrefLabel, '^Microinfarct of heart$')
                || regex(?synonym, '^Microinfarct of heart$'))
      }
}
```

childPrefLabel	synonym	parentPrefLabel
"Microinfarct of heart"@EN	"Microinfarct of heart (disorder)"@EN	"Myocardial infarction"@EN
"Microinfarct of heart"@EN	"Microinfarction of heart"@EN	"Myocardial infarction"@EN

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. FHIR AND SEMANTIC RECONCILIATION OF MEDICAL CONCEPTS

The FHIR standard for exchanging healthcare information electronically has been touted as the lingua franca that would help resolve the existing heterogeneity in healthcare data formats which could then be leveraged to reconcile the underlying semantics. FHIR is also intended to support health information exchange and provide a means for mHealth applications to easily gather data from multiple HIT systems (EMRs/EHRs/PHRs) via the common FHIR standard. While attaining the first goal can be trivial if all the stakeholders in the healthcare field adopt the same data exchange format, namely FHIR, there is still the problem of ensuring that the semantics of the exchanged data is preserved from one record to another. To achieve semantic interoperability, FHIR advocates that the data be encoded using standardized coding schemes/systems (e.g., SNOMED-CT for problems, LOINC for observations, RxNorm for medications, etc.), and that the text to display be one of the preferred names. However, nothing in the standard prevents the use of two different coding systems for the same issue (e.g., SNOMED-CT and ICD-10-CM for medical problems). Consider, for example, the two FHIR records in Figures 16 and 17, respectively (abridged for clarity), the first from a physician in a walk-in clinic using SNOMED-CT to report a medical condition (Figure 16), and the second from a family doctor using ICD-10-CM to record the same medical condition, namely a common cold (Figure 17).

Despite the fact that both records in Figures 16 and 17 comply with the same exchange standard (FHIR), it is impossible for the system and/or machine to infer from the provided information that both records actually describe the same medical condition. Whereas, by using the approach we presented in Section 4.2 and 4.3, we can quickly obtain RDF triples from which to infer the semantic equivalence of the two medical concepts. Due to space constraints, we limit the presentation to two quick queries evidencing that both medical conditions share the same semantics. The first query for SNOMED with SPARQL (Figure 18) and its results (Figure 19) clearly shows that the reported condition ("Common cold") is also known as an "Acute nasopharyngitis" which is the medical term used in the second FHIR/XML record in Figure 21.

Figure 16. FHIR/XML message using SNOMED-CT to report a medical condition

```
<Condition xmlns="http://hl7.org/fhir">
  <id value=" f002"/>
  <!-- abridged for clarity purposes -->
  <patient>
    <reference value="Patient/f002"/>
    <display value= "John Doe"/>
  </patient>
  <code>
    <coding>
      <system value="http://snomed.info/sct"/>
      <code value="82272006"/>
      <display value=" Common cold "/>
    </coding>
  </code>
  <!-- abridged for clarity purposes -->
</Condition>
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 17. FHIR/XML message using ICD-10-CM to report the same medical condition

```
<Condition xmlns="http://hl7.org/fhir">
  <id value=" f005"/>
  <!-- abridged for clarity purposes -->
  <patient>
    <reference value="Patient/f005"/>
    <display value= "John Doe"/>
  </patient>
  <code>
    <coding>
      <system value="http://hl7.org/fhir/sid/icd-10-us"/>
      <code value="J00"/>
      <display value=" Acute nasopharyngitis "/>
    </coding>
  </code>
  <!-- abridged for clarity purposes -->
</Condition>
```

Figure 18. SPARQL Query and for SNOMED-CT

```
PREFIX skos:  <http://www.w3.org/2004/02/skos/core#>
PREFIX sno:   <http://purl.bioontology.org/ontology/SNOMEDCT/>
SELECT DISTINCT  *
WHERE {
          { sno:82272006  skos:prefLabel  ?prefLabel . }
          UNION
          { sno:82272006  skos:altLabel  ?synonym . }
 }
```

Figure 19. SPARQL results for SNOMED-CT

prefLabel	synonym
"Common cold"	--
--	"Acute rhinitis"
--	"Acute coryza"
--	"Common cold (disorder)"
--	"Infective rhinitis"
--	"Acute nasal catarrh"
--	"Acute nasopharyngitis"
--	"Acute infective rhinitis"
--	"Head cold"
--	"Infective nasopharyngitis"
--	"Cold"
--	"Acute nasopharyngitis, NOS"
--	"Infective nasopharyngitis, NOS"

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The second query for ICD-10 with SPARQL (Figure 20) and results (Figure 21) confirms that the reported medical condition is a common cold. As already explained, by generating the related triples (using the CONSTRUCT keyword) and storing them locally, we can make the appropriate inferences.

Through this simple example, we have shown how our approach applies to FHIR records to evidence synonymy of medical terms.

6. CONCLUSION

This paper has tackled a quite common issue in attempting to integrate data from multiple EMRs/EHRs into a single PHR in which the co-existence of several records in different systems utilize different medical terms to refer to the same medical problem. To set the stage for the work, Section 2 reviewed background information on different types of health information technology systems (PHRs, EMRs, EHRs) and the role of health information exchange. This was followed in Section 3 by a review of the nomenclature in the medical field and the introduction of a use case scenario of healthcare that was utilized for examples throughout the rest of the paper. Section 4 presented the main focus of our paper, a proposed approach to identify and reconcile synonymy of diverse medical terms occurring in multiple EMRs/EHRs. Our approach in Section 4 queried the relevant coding systems to single out and generate a set of RDF triples asserting known and published synonymy, and inferring from those triples potential semantic equivalences in our medical terms. Section 5 extended the work of Section 4 by exploring the semantic harmonization issue that involves the ability to achieve the desired reconciliation utilizing the new FHIR (2014) standard.

Figure 20. SPARQL query for ICD-10

```
PREFIX  skos:  <http://www.w3.org/2004/02/skos/core#>
PREFIX  icd10: <http://purl.bioontology.org/ontology/ICD10CM/>
SELECT DISTINCT  *
WHERE {
        { icd10:J00  skos:prefLabel  ?prefLabel . }
        UNION
        { icd10:J00  skos:altLabel  ?synonym . }
}
```

Figure 21. SPARQL results for ICD-10

prefLabel	synonym
"Acute nasopharyngitis [common cold]"	--
--	"Acute rhinitis"
--	"Infective rhinitis"
--	"Infective nasopharyngitis NOS"
--	"Nasopharyngitis NOS"
--	"Coryza (acute)"
--	"Nasal catarrh, acute"

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

One final important conclusion is that our approach does not require any specific knowledge of the internal structure of the medical coding systems, and relies on a simple and generic formulation to generate the RDF triples from which further inferred synonymy can be evidenced to the system hosting the PHRs. Note also that for clarity, we have not tried to take advantage of advanced features of SPARQL (BIND, VALUES, sub-SELECT[20], etc.) to optimize or beautify our code[21] (e.g., avoid hard-coded medical strings in queries). An obvious extension to our work would be to implement the presented solution into an open-source PHR system to allow data collected from multiple EHRs/EMRs to be semantically reconciled to be presented to a patient via an mHealth application. This work requires additional algorithms and processes to parse the PHR data to identify the coding systems used and to isolate the medical terms whose semantic equivalence we want to verify. Further, the SPARQL queries would need to be transmitted to the endpoint using a RESTFul protocol. Note also that the work in this paper has been influenced in part by this research on security for digital healthcare (Demurjian, et al. 2014/2016) that includes the attainment of role-based access control in XML for the HL7 standard (De La Rosa Algarin, et al. 2013/2016). Semantic reconciliation will play a vital role in the security of healthcare data to insure that the correct terms are delivered to the authorized individuals.

REFERENCES

Allemang, D., & Hendler, J. (2011). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL* (2nd ed.). Morgan Kaufmann.

Bender, D., & Sartipi, K. (2013). HL7 FHIR: An Agile and RESTful approach to healthcare information exchange. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems* (pp. 326-331). IEEE. 10.1109/CBMS.2013.6627810

Bohring, H., & Auer, S. (2005). Mapping XML to OWL Ontologies. *Leipziger Informatik-Tage*, *72*, 147–156.

Breitling, F. (2009). A standard transformation from XML to RDF via XSLT. *Astronomical Notes - Astronomische Nachrichten, 330* (7) CD-10-CM. (2016). *International Statistical Classification of Diseases, 10th Revision Clinical Modification*. CDC. Retrieved from http://www.cdc.gov/nchs/icd/icd10cm.htm

CDA. (2005). Health Level Seven Clinical document. *Architecture (Washington, D.C.)*, *HL7*, v2. Retrieved from http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

Daglish, D., & Archer, N. (2009). Electronic personal health record systems: a brief review of privacy, security, and architectural issues. *Proceedings of the World Congress on Privacy, Security, Trust and the Management of e-Business CONGRESS'09* (pp. 110-120). IEEE.

De La Rosa Algarin, A., & Demurjian, S. (in press). Extending the UML Standards to Model Tree-Structured Data and their Access Control Requirement. *Proc. of SSR 2016: Security Standardisation Research*, NIST, Gaithersburg, MD, USA. 10.1007/978-3-319-49100-4_8

Demurjian, S., De La Rosa Algarin, A., Bi, J., Berhe, S., Agresta, T., Wang, X., & Blechner, M. (2014/2016). A Viewpoint of Security for Digital Health Care: What's There? What Works? What's Needed? International Journal of Privacy and Health Information Management, 2(1), 1-21.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., & Laleci, G. B. (2005). A survey and analysis of electronic healthcare record standards. *ACM Computing Surveys*, *37*(4), 277–315. doi:10.1145/1118890.1118891

FHIR. (2014). *Health Level seven Fast Healthcare Interoperability Resources*. Retrieved from https://www.hl7.org/fhir/index.html

Genitsaridi, I., Kondylakis, H., Koumakis, L., Marias, K., & Tsiknakis, M. (2015). Evaluation of personal health record systems through the lenses of EC research projects. *Computers in Biology and Medicine*, *59*, 175–185. doi:10.1016/j.compbiomed.2013.11.004 PMID:24315661

HL7 v2. (1989). *Health Level Seven Version 2 Product Suite*. Retrieved from http://www.hl7.org/implement/standards/product_brief.cfm?product_id=185

Halpin, H., Hayes, P. J., McCusker, J. P., McGuinness, D. L., & Thompson, H. S. (2010). When owl:sameAs isn't the same: An analysis of identity in linked data. *Proceedings of the International Semantic Web Conference* (pp. 305-320). 10.1007/978-3-642-17746-0_20

Health, I. T. (2014). Office of the National Coordinator for Health Information Technology. *HealthIT*. Retrieved from https://www.healthit.gov/providers-professionals/electronic-medical-records-emr

Health, I. T. (2016). Office of the National Coordinator for Health Information Technology. *HealthIT*. Retrieved from https://www.healthit.gov/patients-families/blue-button/about-blue-button

Kawazoe, Y., T Imai, T., & Ohe, K. (2016). "A Querying Method over RDF-ized Health Level Seven v2. 5 Messages Using Life Science Knowledge Resources". *JMIR Medical Informatics, 2016*.

Miles, A., Matthews, B., Wilson, M., & Brickley, D. (2005). SKOS core: simple knowledge organisation for the web. *Proceedings of the International Conference on Dublin Core and Metadata Applications*.

Noy N.F., Shah N.H., Whetzel P.L., Dai B., Dorf M, Griffith N, Jonquet C, Rubin D.L., Storey M.A., Chute C.G., & Musen M.A. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res., 37(Web Server issue). PMCID: PMC2703982

OBO DOID. (2016). *The Human Disease Ontology*. The OBO Foundry. Retrieved from http://www.obofoundry.org/ontology/doid.html

Prasser, F., Kohlmayer, F., Kemper, A., & Kuhn, K. (2012). A Generic Transformation of HL7 Messages into the Resource Description Framework Data Model. *Proceedings of GI-Jahrestagung*; Germany (pp. 1559–1564).

Sahoo, S. S., Halb, W., Hellmann, S., Idehen, K., Thibodeau Jr, T., Auer, S., Sequeda, J., & Ezzat, A. (2009). A survey of current approaches for mapping of relational databases to RDF. W3C RDB2RDF Incubator Group Report.

Salvadores, M., Horridge, M., Alexander, P., Fergerson, R., Musen, M., & Noy, N. (2012). Using SPARQL to query bioportal ontologies and metadata. *The Semantic Web–ISWC*, *2012*, 180–195.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Smits, M., Kramer, E., Harthoorn, M., & Cornet, R. (2015). A comparison of two Detailed Clinical Model representations: FHIR and CDA. *European Journal for Biomedical Informatics*, *11*(2).

SNOMED-CT. (2016*). Systematized Nomenclature of Medicine - Clinical Terms.* Retrieved from http://www.ihtsdo.org/snomed-ct

ENDNOTES

1. "How mHealth is Changing Health and Healthcare": http://www.himss.org/ResourceLibrary/mHimssRoadmapLanding.aspx?ItemNumber=30562
2. By extension, PHRs designate also the computer applications that manage these electronic records
3. "The mHealth Summit: Local & Global Converge": http://caroltorgan.com/mhealth-summit/
4. These two concepts are defined in the next section
5. http://loinc.org/
6. http://www.ihtsdo.org/snomed-ct/
7. http://www.nlm.nih.gov/research/umls/
8. https://www.healthit.gov/policy-researchers-implementers/about-onc-health-it-certification-program
9. Himss.org, 'Meaningful use stage 3 final rule', 2015. [Online]. Available: http://www.himss.org/ResourceLibrary/genResourceDetailPDF.aspx?ItemNumber=44987 [Accessed: 11- May- 2016]
10. https://www.w3.org/2004/02/skos/
11. https://www.w3.org/RDF/
12. https://www.w3.org/TR/rdf-sparql-query/
13. https://www.hl7.org/fhir/rdf.html
14. Found at: any23.org/
15. http://www.bioontology.org/BioPortal
16. https://www.w3.org/2004/02/skos/
17. We discarded skos:exactMatch and skos:closeMatch because reminiscent of mechanical string comparison
18. The search should only consider active concepts in the ontology (CONCEPTSTATUS = 0)
19. Our data had to be slightly edited to comply with Protégé's requirements
20. Some of these features don't seem to be supported by the BioPortal SPARQL endpoint
21. The presented code has been validated and executed via the demo BioPortal endpoint.

This research was previously published in the International Journal of Information Technology and Web Engineering (IJITWE), 12(2); edited by Ghazi I. Alkhatib; pages 26-48; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Message in Figure 1 Translated into FHIR/RDF

```
@prefix fhir: <http://hl7.org/fhir/> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix sct: <http://snomed.info/id/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://hl7.org/fhir/Conditionf001> a fhir:Condition;
  fhir:nodeRole fhir:treeRoot;
  fhir:Resource.id [ fhir:value "f003"];
  fhir:DomainResource.text [
     fhir:Narrative.status [ fhir:value "generated" ];
     fhir:Narrative.div "<div xmlns=\"http://www.w3.org/1999/xhtml\">
Myocardial infarction, June 15, 2016</div>"
  ];
  fhir:Condition.patient [
     fhir:link <http://hl7.org/fhir/Patient/f001>;
     fhir:Reference.reference [ fhir:value "Patient/f001" ];
     fhir:Reference.display [ fhir:value "Mr. John Smith" ]
  ];
  fhir:Condition.code [
     fhir:CodeableConcept.coding [
       fhir:index 0;
       fhir:concept sct:22298006;
       fhir:Coding.system [ fhir:value "http://snomed.info/sct" ];
       fhir:Coding.code [ fhir:value "22298006" ];
       fhir:Coding.display [ fhir:value "Myocardial infarction" ]
     ]
  ];
  fhir:Condition.category [
     fhir:CodeableConcept.coding [
       fhir:index 0;
       fhir:concept sct:439401001;
       fhir:Coding.system [ fhir:value "http://snomed.info/sct" ];
       fhir:Coding.code [ fhir:value "439401001" ];
       fhir:Coding.display [ fhir:value "diagnosis" ]
     ]
  ];
  fhir:Condition.clinicalStatus [ fhir:value "active"];
  fhir:Condition.verificationStatus [ fhir:value "confirmed"];
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
fhir:Condition.severity [
   fhir:CodeableConcept.coding [
     fhir:index 0;
     fhir:concept sct:6736007;
     fhir:Coding.system [ fhir:value "http://snomed.info/sct" ];
     fhir:Coding.code [ fhir:value "6736007" ];
     fhir:Coding.display [ fhir:value "Moderate" ]
   ]
];
fhir:Condition.onsetDateTime [ fhir:value "2011-08-05"^^xsd:date].
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 30
Principles and Analysis of Biological Networks:
Biological Pathways and Network Motifs

Manika Sehgal
Institute of Microbial Technology, India

TirathaRaj Singh
Jaypee University of Information Technology, India

ABSTRACT

The biological network complexity is growing enormously and in order to reveal confined properties of these intricate networks, detection of crucial network components may assist in gaining effortless perceptive on the underlying biological processes. Analyzing complex biological pathways for their disease association is still a drawn-out process and requires an integrative approach for comprehensive examination of proteins and interactions to identify candidate markers underlying major malignancies and genetic disorders. There is a need for an amalgamated approach to annotate all the sub-components and their associated interactions in a biological system. It is anticipated that analysis of biological pathways would serve as a valuable accompaniment for analyzing biomarkers in disease pathways and will also contribute scientific knowledge towards their better understanding.

INTRODUCTION

In present era, the advent of human genome sequence, high-throughput techniques and massive generated information has revolutionized the biomedical research (Bentley D. R., 2006). This huge amalgamation of genomic data is rising exponentially due to employment of next generation technologies and application of modern biology (Bentley D. R. 2006; Voelkerding. K. V et al., 2009; Mamanova L, et al. 2010). However, less emphasis is remunerated on the obligatory genomic analysis; subsequently there is a need to apply extensive computational approaches that could deal abreast with the remarkable

DOI: 10.4018/978-1-7998-1204-3.ch030

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

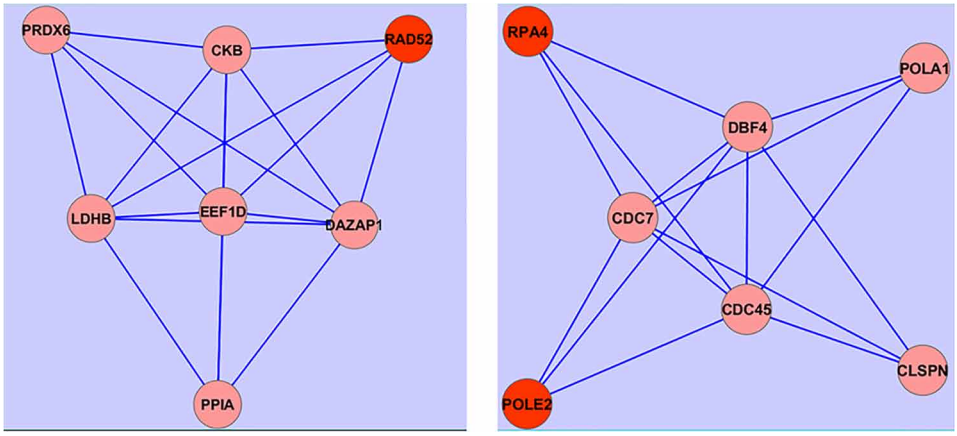
growth in generated biological data (Barabasi A. L. & Oltvai Z.N. 2004; Jeong H., et al. 2000, Jeong H. et al. 2001). Since, the human system is very complex and the ultimate challenge is to decipher numerous multifaceted processes involved in its regulation therefore biological network analysis has major relevance which assists in wide-ranging understanding of an entire system as a whole (Jeong H., et al. 2001). These biological networks present several processes such as cellular mechanisms, signaling cascades, transcriptional regulation, expression profiles and protein-protein interactions (Barabasi A. L. & Oltvai Z. N. 2004; Jeong H. et al., 2000). In Figure 1, these networks comprise of nodes as entities and edges as relationships among them. In general terms a network may be any social interaction, electric circuits, internet and food chain, etc. In biological context, these nodes may vary from a range of macromolecules for instance DNA, RNA, proteins and metabolites. The applications of networks can be observed in ecological networks, expression networks, gene regulatory networks, metabolic networks, protein interaction networks.

Recently, pathway analysis and reconstruction has also gained a lot of attention as it assists in unravelling unknown parameters concerning many precarious diseases. Therefore, the study of networks, their modelling, visualization and simulation are an important aspect for extensively understanding the underlying processes to make sense of the complex available data. Biological networks can be broadly classified into following types:

1. Protein-Protein Interaction (PPI) Networks

The PPI networks form the basis for a vast majority of cellular events, including signal transduction and transcriptional regulation. An extensive understanding of a biological system is feasible only once all the interactions are characterized. These interactions are very well deduced from numerous high-throughput experiments such as protein arrays, co-immunoprecipitation, microarrays and predicted from computational resources developed over time. These *in silico* tools and databases include Search Tool for the Retrieval of Interacting Genes/proteins (STRING), Biomolecular Interaction Network Database (BIND), Molecular INTeraction database (MINT) and Database of Interacting Proteins (DIP). These PPI databases help us effortlessly access the existing information for biological interactions among a

Figure 1. Representation of a sub-network comprising of nodes as entities and edges as relationship/ association with nodes



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

variety of species. The published literature and methods used to exemplify the interaction studies are also described. These resources boost up the understanding of underlying processes which aid the scientific community for designing their future experiments.

2. Signal Transduction and Gene Regulatory Networks

The gene regulatory networks consist of clusters of genes (or gene products) with evidence of co-expression and are generally associated with cascade of processes. In this form of networks, connections usually represent degrees of co-expression where the in-depth knowledge of process is not necessary. Here, the networks are usually non-predictive in nature. The signal transduction networks are induced from some external signals (neurotransmitters, peptides, growth factors) followed by a series of processes involving binding with receptors.

3. Expression Networks

The expression networks deal with the sequence of changes occurring from transcription of DNA into mRNA which is the most regulated step in gene expression. Further, mRNA encodes for proteins i.e. directly proportional to the expression status. Most genes are regulated at the transcriptional level and 5-10% of protein-encoding genes encode regulatory proteins. The interactions between regulated genes and regulatory proteins constitute a complex web called transcriptional network or expression network.

4. Metabolic Networks

The metabolic networks comprises of all the interactions that sums up for the entire chemical reactions taking place within a cell. These processes are implicated mainly in providing energy for the vital cellular and metabolic processes. The metabolic network represents a series of chained reactions where the connections depict quantifiable contacts between the molecules. Since, a metabolic network illustrates a chained reaction, the enzymatic process is evidently elucidated where the changes are predictable downstream.

Currently, the spotlight of major resources is on the metabolic pathways, signaling and regulatory networks to fill in the research gaps and missing links between pathways. Ultimately, all these networks comprising of signal transduction, gene regulatory, metabolic and protein-protein interactions further interact with each other and build a complicated system where these networks are species–specific and may vary among different taxas (Mamanova L, et al. 2010, Jeong H., et al. 2001*)*. There are a variety of macromolecules like enzymes, transcriptional regulators, ribosomes, other biological entities that form complex biological networks to carry out fundamental cellular functions.

BACKGROUND

For a comprehensive perspective on biological complexity and underlying intricate networks, exclusive monitoring of whole system to component level entities is required. Initially, entire pathway is analyzed which is a sequence of interactions between biochemical compounds for controlling the flow of information and energy in the cell. There are myriad of molecular compounds, molecules and cofactors that are

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

component of a network. These include proteins, their complexes, amino acids, peptides, sugars, lipids, cofactors, prosthetic groups, metabolites and other essential compounds like Sulphur, phosphorous, iron and radicals. The analysis of biochemical pathways have become easier with the advancements in experimental technologies like microarrays, proteome analysis, yeast two-hybrid system and phosphorylations of proteins which deal with understanding of interactions among entities. Once, the PPI interactions are clearly elucidated, the reconstruction of a biological pathway becomes simpler which may further be subjected to simulation for estimating the system dynamics. This domain of pathway reconstruction and classification has led to an emerging field of network analysis that assists in wide-ranging analysis of biological pathways. This area basically employs bioinformatics approaches and utilizes computational tools for carrying out functional enrichment, annotation and defining ontologies for a biological system. The results from pathway/network analysis can also be subjected to gene ontology (GO) analysis for governing their involvement in biological processes, cellular compartments and molecular functions. The key feature for network analysis is to postulate vital conserved patterns that may be functionally significant which in turn reveals candidate genes/proteins.

The complexity of the genome becomes multi-folded as the pathways consecutively chain together to form larger pathways and intricate networks. Specific conserved patterns in these networks contribute to the disease etiology in several species. Although, cancers and multi-system defects are a major threat to mankind but still dealing with harmful consequences of these damages is not adequately feasible and also there is no absolute therapy for cancer. The understanding of major cellular processes and its implications through network analyses could help reveal important aspects regarding the same. There is an imperative need to understand these biological mechanisms and triggered signalling cascades so that damages could be sensed well in advance. Upon sensing the damage, factors involved in the regulation of pathways should be clearly elucidated along with their biological roles and positions where they exert their functions (Jeong H., et al. 2001, Jeong H., et al. 2000). Thus, to unravel the complexity hidden in intricate biological phenomenon, a pathway-level analysis for biological networks is essential for the interaction of genes/proteins/networks for understanding myriad of cellular processes.

The knowledge pertaining to interacting partners assists in examining the impact of individual proteins in the overall pathway. In general, a network has a highly complex structure comprising of thousands of nodes with connected edges representing general properties. In order to estimate these properties, absolute acquaintance on kinetics and functional dynamics is obligatory. Once, the node, edges and their properties are calibrated the ultimate challenge is to bridge both the levels. Although, the redundancy and conserved patterns can easily be exemplified further leading to identification of candidate genes/proteins in disease specific pathways. Additionally, analyzing the interaction architecture even helps in setting up a framework for mathematic models where analyzing individual components may not provide meaningful insights; therefore whole pathway analysis is the primary requisite for the study of complex systems.

MAIN FOCUS OF THE CHAPTER

In recent times, various bioinformatics and statistical approaches have been applied to thoroughly understand the structural, functional and evolutionary evidences latent in biological networks of interest. Numerous approaches are available to elucidate important findings from these intricate biological pathways, their comparisons, reconstructions and designing. Figure 2 represents the methodology that could be applied for identifying key players from a biological network. Additionally, vital network components

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

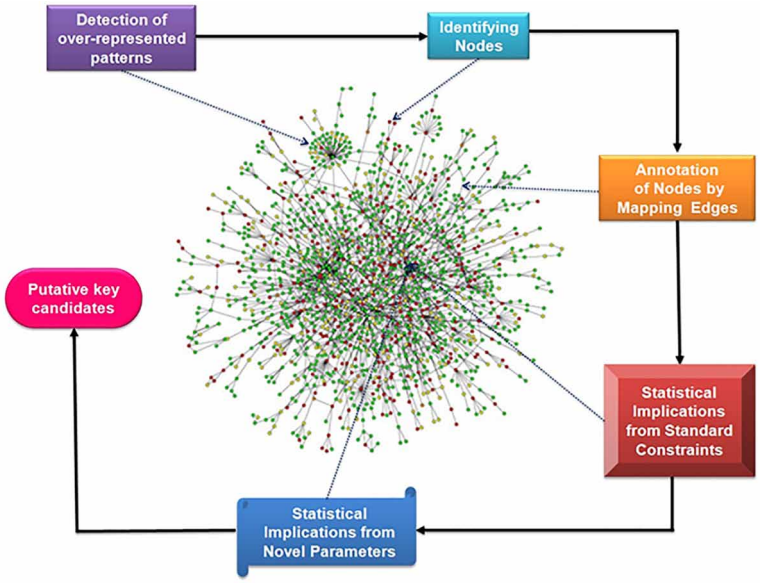
referred as network motifs (Alon U, 2007), found in elevated frequencies that could generally be expected by chance in a pathway could also be identified, as in Figure 3. These network motifs provide statistically overrepresented sub-structures (sub-graphs) in a network and are recognized as simple building blocks of a complicated network. These network motifs play a central role in recognition and analysis of specific patterns in biological networks and yield significant insights into understanding complex biological processes involved in intricate human diseases (Kim W, et al. 2011). Major key candidates and network components or motifs can also be deduced from these complex interactions therefore confining the search to a few important genes.

To thoroughly understand the dynamics of a biological system, a comprehensive perspective on the complex interactions and its topology is required. The chapter highlights details on network properties (Assenov Y, et al. 2008) such as shortest path length, clustering coefficient, neighbourhood connectivity and other crucial parameters as depicted in Figure 4. The shortest path length $L(n,m)$ is the length of shortest path between two nodes n and m. The shortest path length distribution gives the number of node pairs (n,m) with $L(n,m) = k$ for $k = 1,2,....,z$. Whereas, the neighborhood connectivity of a node n is defined as the average connectivity of all its neighbours. The neighborhood connectivity distribution gives the average of the neighborhood connectivity's of all nodes n with k neighbors for $k = 0,1,....,z$. The average clustering coefficient distribution gives the average of clustering coefficients for all nodes n with k neighbors for $k = 2,...,z$. In directed network, the clustering coefficient C_n of a node n is defined as:

$$C_n = e_n/(k_n(k_n\text{-}1))$$

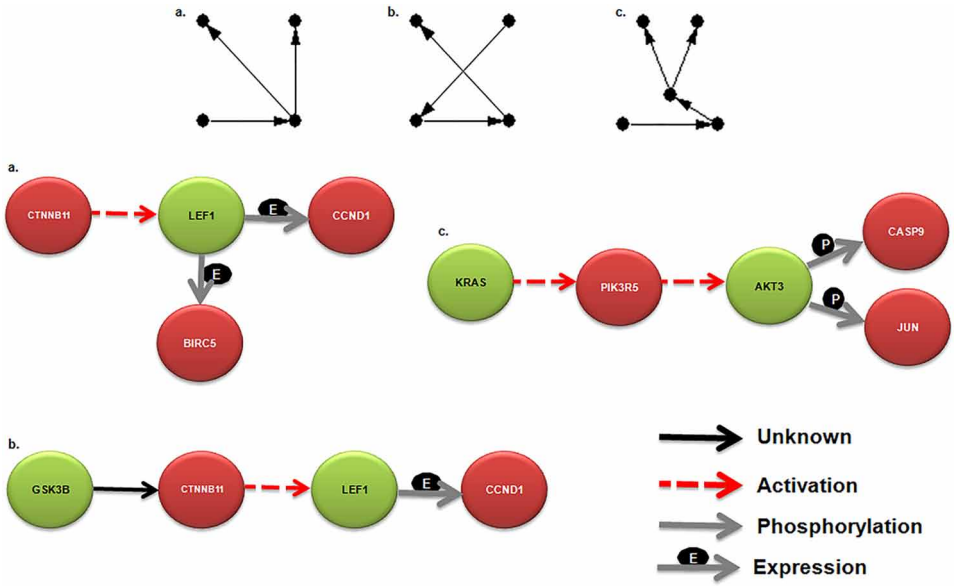
where k_n is the number of neighbors of n and e_n is the number of connected pairs between all neighbors of n. The clustering coefficient is a ratio N / M, where N is the number of edges between the neighbors

Figure 2. Methodology for detecting key candidates from the pathway



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Identified and annotated network motifs in colorectal cancer
Sehgal M., et al., 2015.



of *n*, and *M* is the maximum number of edges that could possibly exist between the neighbors of *n*. The clustering coefficient of a node always ranges from 0-1. The average clustering coefficient has already presented its utilization in metabolic networks where the modular organization has been studied.

COMPUTATIONAL METHODS AND APPROACHES AVAILABLE FOR PATHWAY ANALYSIS

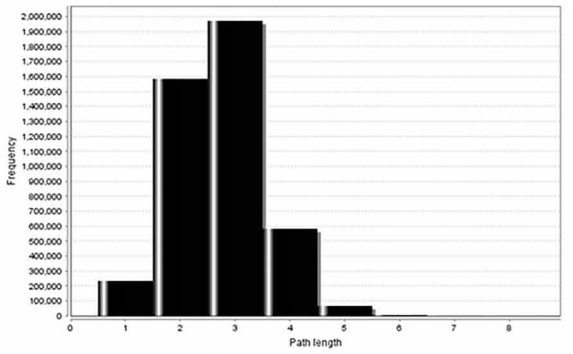
The rapid growth of knowledge on functional analysis of genes, proteins and cellular processes demands wide-ranging databases and novel computational methods and approaches to be devised for efficient representation of proteins as well as their functional interactions whereas the chapter extensively focus on the important available approaches. The methods and resources (Smoot M. E.,et al. 2011, Wernicke S. & Rasche F. 2006, Schreiber F. & Schwobbermeyer H. 2005, Kashtan N.,et al., 2004) for studying biological networks include:

Cytoscape

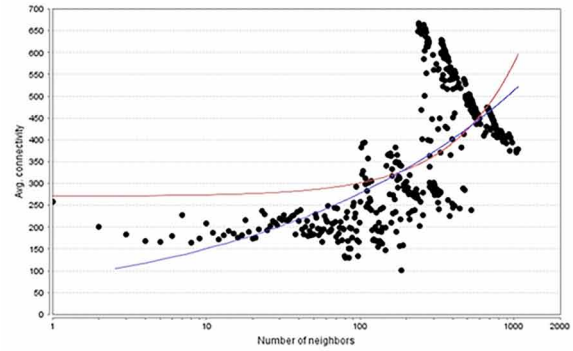
Cytoscape is a very widely used open source software platform for visualization of molecular interaction networks and biological pathways. It allows integration of the networks with single nucleotide polymorphism (SNP) data, phenotypes, gene expression profiles and other molecular states leading to annotations and functional enrichment. Cytoscape has an additional feature for loading nodes, edges and networks where different attributes could be mapped for extensive understanding of biological processes. User can easily access and analyze the expression results and GO annotations in the platform. There are various apps and plugins available for annotation of these networks incorporating data from a variety of datasets

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

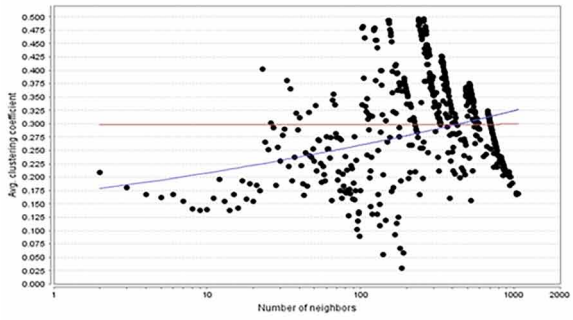
Figure 4. Statistical parameters elucidating significance of the biological pathway



a. Shortest path length



b. Neighborhood connectivity (In and Out)



c. Average clustering coefficient

from microarrays and high-throughput studies. These plugins are generally freely available and stimulate easy visualization of complex biological data. A classical example of statistical inferences made from network analyzer plugin has already been observed in Figure 4. It has numerous features available for data integration, analysis, visualization, molecular profiling analyses, new layouts, additional file formats support and connection with databases. Cytoscape supports many standard network and annotation file formats such as SIF (Simple Interaction Format), BioPAX, PSI-MI, GraphML, KGML (KEGG XML), SBML and Gene Association. Cytoscape is considered most influential when utilized in juxtaposition with larger databases of protein-protein, protein-DNA and genetic interactions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fanmod

Fanmod is a tool for fast network motif detection. It is better than its counterparts (explained below) in terms of efficiency and speed when dealing with larger motifs generally more than 8-node sub- graphs. It randomly detects over-represented sub-graphs in a network called as network motifs which are recurring patterns of interactions and are statistically significant. It exploits the most efficient RAND-ESU algorithm for the detection of motifs. It provides with motifs ranging from 3-8 node sub-graphs with colored edges and have filtering done on the basis of z-score and p-values. Fanmod can be easily downloaded on a system and be used as a standalone software which performs calculations at a faster rate.

MAVisto

Like Fanmod, MAVisto is used for the exploration of motifs in a network. It provides a flexible motif search algorithm and different views for the analysis and visualization of network motifs. The MAVisto tool is written in Java and has an editor for graphs visualization and a toolkit for implementing graph algorithms. It is also available freely for academic purposes at http://mavisto.ipk-gatersleben.de/mavisto.jnlp. Usually, three different types of analyses can be performed with MAVisto by considering only edge labels, only vertex labels or both, edge and vertex labels.

Mfinder

The motif finder (mfinder) is another software tool for detecting network motifs. It identifies the network motifs in biological pathways by using two algorithms i.e. firstly it takes into account the full enumeration of sub-graphs and then sampling of sub-graphs is done for estimating the sub-graph concentrations. It employs switching method for the generation of random motifs. The standalone version can be easily downloaded from the following URL

https://www.weizmann.ac.il/mcb/UriAlon/download/network-motif-software.

The basic steps for the identification of network motifs from any of the above mentioned tool comprise of the following basic steps:

1. Identify all sub-graphs of n nodes in the network.
2. Randomize the network, while keeping the number of nodes, edges, and degree distribution unchanged.
3. Identify all sub-graphs of n nodes in the randomized version.
4. Sub-graphs that occur significantly more frequently in the real network, as compared to the randomized one, are designated as motifs.

RESOURCES AND DATABASES FOR BIOLOGICAL NETWORKS

The genome sequencing and the latest research strategies have generated huge and complex genomic and proteomic data which is available in repertoire of online resources and databases. Enormous research

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tools are also available for the extensive analysis of these complex interactomes. There are a variety of resources on biological networks and their analysis available online which can be classified on different means as follows:

General Pathway Repositories

This module supplies information on available general biological pathway databases. The pathway databases mainly highlight entire cascade of processes instead of individual components. Although it do provide insights into smaller components or interacting proteins but the ultimate aim is to understand the system as a whole.

BioCyc/MetaCyc

It stands for biological encyclopedia/ metabolic encyclopedia. Biocyc is a pathway/genome database for organisms with completely sequenced genomes where species-specific pathways are inferred form MetaCyc whereas MetaCyc is a non-redundant reference database for metabolic pathways, reactions, enzymes and compounds.

Kyoto Encyclopedia of Genes and Genomes (KEGG)

KEGG is mainly considered as a common platform for consolidated set of databases that cover genomics (GENE), chemical compounds (LIGAND) and reaction networks (PATHWAY). It has a broad coverage related to other databases concerning on metabolics, signal transduction, disease, *etc*. It contains pathways from numerous species and allows only static visualization services where the pathways can be downloaded in KGML format for further annotation and analysis (Kanehisa, M.,et al. 2016).

BioCarta

It is a corporate-owned and publicly curated pathway database which provides an interactive user interface with cartoon pathway maps. BioCarta focuses mainly on human and mouse related biological pathways.

BioModels Database

BioModels is a database on published quantitative models for biochemical processes. It contains information on all models/pathways that are curated manually. It allows the models to be created in SBML format for quantitative modeling.

Some other pathway databases include PathDB, Signaling PAthway Database (SPAD) and Cytokine Signaling Pathway Database.

Protein Interaction Databases

The protein interaction databases section provides an exclusive compendium for various species-specific, function-specific, interaction type-specific and general protein-protein interaction databases.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. **Species-Specific Databases:**
   a. **FlyNets:** Gene networks in the fruit fly,
   b. **MIPS:** Yeast Genome Database,
   c. **RegulonDB:** A DataBase On Transcriptional Regulation in E. Coli, and
   d. **PIMdb:** Drosophila Protein Interaction Map database.
2. **Function-Specific Databases:**
   a. Biocatalysis/Biodegradation Database,
   b. **BRITE:** Biomolecular Relations in Information Transmission and Expression,
   c. **COPE:** Cytokines Online Pathfinder Encyclopaedia,
   d. Dynamic Signaling Maps,
   e. **EMP:** The Enzymology Database,
   f. **FIMM:** A Database of Functional Molecular Immunology, and
   g. **CSNDB:** Cell Signaling Networks Database.
3. **Interaction Type-Specific Databases:**
   a. **DIP:** Database of Interacting Proteins,
   b. **DPInteract:** DNA-protein interactions,
   c. **Inter-Chain Beta-Sheets (ICBS):** A database of protein-protein interactions mediated by interchain beta-sheet formation,
   d. **Interact:** A Protein-Protein Interaction database, and
   e. GeneNet (Gene networks).
4. **General Protein-Protein Interaction Databases:**
   a. **BIND:** Biomolecular Interaction Network Database,
   b. **BindingDB:** The Binding Database,
   c. **MINT:** A database of Molecular INTeractions,
   d. **PATIKA:** Pathway Analysis Tool for Integration and Knowledge Acquisition,
   e. **PFBP:** Protein Function and Biochemical Pathways Project, and
   f. PIM (Protein Interaction Map) .

PATHWAY EXCHANGE FORMATS

There are numerous pathway exchange formats developed for easy exchange of information from one platform to another. The exchange formats help maintain a standard for representing pathway data or easy understanding among different researchers. Some of the exchange formats have been described below:

Extensible Markup Language (XML)

XML is a standard of representing information in a machine-readable manner. It is a division of SGML and is much better than hyper text markup language (HTML) in terms of its extensibility in defining tags that enclose data. For instance:

```
<XMLInfo>
        <someTag>Some data here</someTag>
        <anotherTag>Additional data </anotherTag>
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
        <attributeTag data="embedded in tag" />
</XMLInfo>
```

Systems Biology Markup Language (SBML)

SBML is a representation format based on XML for exchange and storage of computational models of biological processes. It is a machine-readable exchange format and its strength lies in the representation phenomena at the scale of biochemical reactions. Several tools use this format for standard input and output operations in order to remove chances for errors in translation and assuring a common starting point for analyses and simulations *(Keating, S. M, et al. 2006).* SBML's framework is suitable for representing models commonly found in research on topics like gene regulation, cell signaling pathways, metabolic pathways, biochemical reactions, and many more. Example of its format is given below:

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns=http://www.sbml.org/sbml/level1
level="1" version="2">
<model name="gene_network_model">
  <listOfUnitDefinitions>
        . . .
  </listOfUnitDefinitions>
<listOfCompartments>
        . . .
</listOfCompartments>
<listOfSpecies>
        . . .
</listOfSpecies>
<listOfParameters>
        . . .
</listOfParameters>
<listOfRules>
        . . .
</listOfRules>
<listOfReactions>
        . . .
</listOfReactions>
   </model>
</sbml>
```

ISSUES, CONTROVERSIES, PROBLEMS

The pathway analysis as a whole is a cumbersome task and there has been many controversies on its analysis as some researchers suggest studying individual interactions can be more beneficial instead being spellbound by the complicated system whereas others support the idea of systems biology as the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

biological systems are basically dynamic in nature. In the knowledge of intricate pathways/networks, regulatory DNA motifs are considered as bar codes to retrieve the building blocks which may help in gaining better insights in pathway studies. A combination of high-throughput experiments, prior knowledge and bayesian network inference is also indispensable for thorough understanding of a biological problem. Furthermore, efficient analyses on biological networks require robust computational infrastructure for specialized housing, accessing and analyses of ever-growing and amplifying complex biological data. Usually, the complete network of integrated biological pathways is the blueprint of life whereas the genome is only an archiving system of building blocks. Therefore, it is proposed that a better idea is to follow an integrative approach where both the component-levels and system-levels are taken into consideration. In order to achieve this goal, top-down and bottom-up techniques could be applied in an amalgamation.

SOLUTIONS AND RECOMMENDATIONS

The biological pathway/network analysis is a hot spot in current research domain and bioinformatics. This area has a variety of sub-fields that could be taken up by young minds to solve the mysteries of various biological problems. The fundamental themes that could be extensively deliberated include development of new algorithms, tools and comprehensive databases for understanding and visualizing pathways with noteworthy statistical inferences. The areas of research focus on valid implications of metabolic pathways from genomes, designing and development of schemas for constructing pathway related databases and classification systems for biological pathway data.

FUTURE RESEARCH DIRECTIONS

Since, major tools and resources dealing with pathway data don't have a common format for exchange of information and results, work primarily on developing exchange formats can be another valuable alternative in this area. Although systems biology markup language (SBML) solves the problem to some extent but for a consistent and straightforward perceptive, other easily understood formats are the need of the hour for exchange of pathway data. Reconstruction of biological pathways via layout algorithms from the knowledge obtained from interaction studies, simulation and molecular dynamics analysis could also yield significant insights in major biological problems. The study of protein-protein interactions has provided important insights into the functions of many of the known oncogenes, tumor suppressors, and DNA repair proteins assisting in cancer research. Moreover, pathway analysis has evidently proved its implication in the domain of pharmacogenetics research targeting specific drug transporters, drug receptors and drug targets. For any research area such as vaccine development, treatments, designing disease strategies, preventing symptoms, clinical trials and drug development to be promising, the major challenge lies in the understanding of disease mechanisms which is feasible only through widespread pathway and system-level understanding.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

The chapter endows with crucial insights on major biological networks, their reconstruction and visualization, computational methods and approaches available for their analyses, online resources and databases containing information on these networks and pathways. The main challenge till date is to find appropriate targets in a diseased pathway leading to drug discovery process which has been a time consuming and expensive affair. Thus, knowledge of biological pathways will assist the readers and scientific community in gaining insights on intricate networks and available tools and resources for their study. The pathway level analysis of these networks may provide vital clues about putative but predictive biomarkers for numerous diseases. Additionally, top-down and bottom-up approaches could be applied for the annotation of complex biological networks and this functional enrichment would generate biological meaningful information. It is expected that once the biological networks are broadly analyzed, it will also provide innovative paradigms for genetic susceptibility, prevention, diagnosis and rational therapy.

REFERENCES

Alon, U. (2007). Network motifs: Theory and experimental approaches. *Nature Reviews. Genetics*, *8*(6), 450–461. doi:10.1038/nrg2102 PMID:17510665

Assenov, Y., Ramirez, F., Schelhorn, S.-E., Lengauer, T., & Albrecht, M. (2008). Computing topological parameters of biological networks. *Bioinformatics (Oxford, England)*, *24*(2), 282–284. doi:10.1093/bioinformatics/btm554 PMID:18006545

Barabasi, A. L., & Oltvai, Z. N. (2004). Network biology: Understanding the cells functional organization. *Nature Reviews. Genetics*, *5*(2), 101–113. doi:10.1038/nrg1272 PMID:14735121

Bentley, D. R. (2006). Whole-genome re-sequencing. *Current Opinion in Genetics & Development*, *16*(6), 545–552. doi:10.1016/j.gde.2006.10.009 PMID:17055251

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C. A., ... Karp, P. D. (2014). The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research*, *42*(1), D459–D471. doi:10.1093/nar/gkt1103 PMID:24225315

Jeong, H., & ... . (2000). The large-scale organization of metabolic networks. *Nature*, *407*(6804), 651–654. doi:10.1038/35036627 PMID:11034217

Jeong, H., Mason, S. P., Barabási, A.-L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature*, *411*(6833), 41–42. doi:10.1038/35075138 PMID:11333967

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*, *44*(D1), D457–D462. doi:10.1093/nar/gkv1070 PMID:26476454

Kashtan, N., Itzkovitz, S., Milo, R., & Alon, U. (2004). Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics (Oxford, England)*, *20*(11), 1746–1758. doi:10.1093/bioinformatics/bth163 PMID:15001476

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Keating, S. M., Bornstein, B. J., Finney, A., & Hucka, M. (2006). SBMLToolbox: An SBML toolbox for MATLAB users. *Bioinformatics (Oxford, England)*, *22*(10), 1275–1277. doi:10.1093/bioinformatics/btl111 PMID:16574696

Kim, W., Li, M., Wang, J., & Pan, Y. (2011). Biological network motif detection and evaluation. *BMC Systems Biology*, *5*(Suppl 3), S5. doi:10.1186/1752-0509-5-S3-S5 PMID:22784624

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., ... Turner, D. J. (2010). Target-enrichment strategies for next- generation sequencing. *Nature Methods*, *7*(2), 111–118. doi:10.1038/nmeth.1419 PMID:20111037

Schreiber, F., & Schwobbermeyer, H. (2005). MAVisto: A tool for the exploration of network motifs. *Bioinformatics (Oxford, England)*, *21*(17), 3572–3574. doi:10.1093/bioinformatics/bti556 PMID:16020473

Sehgal, M., Gupta, R., Moussa, A., & Singh, T. R. (2015). An Integrative Approach for Mapping Differentially Expressed Genes and Network Components Using Novel Parameters to Elucidate Key Regulatory Genes in Colorectal Cancer. *PLoS ONE*, *10*(7), e0133901. doi:10.1371/journal.pone.0133901 PMID:26222778

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics (Oxford, England)*, *27*(3), 431–432. doi:10.1093/bioinformatics/btq675 PMID:21149340

Voelkerding, K. V., Dames, S. A., & Durtschi, J. D. (2009). Next-generation sequencing: From basic research to diagnostics. *Clinical Chemistry*, *55*(4), 641–658. doi:10.1373/clinchem.2008.112789 PMID:19246620

Wernicke, S., & Rasche, F. (2006). FANMOD: A tool for fast network motif detection. *Bioinformatics (Oxford, England)*, *22*(9), 1152–1153. doi:10.1093/bioinformatics/btl038 PMID:16455747

KEY TERMS AND DEFINITIONS

Annotation: Annotation is defined as a measure of estimating the properties of an entity (gene, protein, metabolite, pathways) such as its functions, localizations and implicated processes. Annotation can be done through manual curation or automation depending on the availability of resources.

Network Motifs: Network motifs are defined as over-represented sub-graphs in a network. These are recurrent patterns occurring in a network with higher frequency compared to other randomly generated networks.

This research was previously published in Library and Information Services for Bioinformatics Education and Research edited by Shri Ram; pages 112-129; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 31
Multiple Sequence Alignment Optimization Using Meta-Heuristic Techniques

Mohamed Issa
Zagazig University, Egypt

Aboul Ella Hassanien
Cairo University, Egypt

ABSTRACT

Sequence alignment is a vital process in many biological applications such as Phylogenetic trees construction, DNA fragment assembly and structure/function prediction. Two kinds of alignment are pairwise alignment which align two sequences and Multiple Sequence alignment (MSA) that align sequences more than two. The accurate method of alignment is based on Dynamic Programming (DP) approach which suffering from increasing time exponentially with increasing the length and the number of the aligned sequences. Stochastic or meta-heuristics techniques speed up alignment algorithm but with near optimal alignment accuracy not as that of DP. Hence, This chapter aims to review the recent development of MSA using meta-heuristics algorithms. In addition, two recent techniques are focused in more deep: the first is Fragmented protein sequence alignment using two-layer particle swarm optimization (FTLPSO). The second is Multiple sequence alignment using multi-objective based bacterial foraging optimization algorithm (MO-BFO).

INTRODUCTION

Bioinformatics is a field that combines computer science and mathematics for analyzing and managing biological data. Developing large databases and complex tools for gene and protein analysis and modeling are the main tasks of bioinformatics besides organization, storing and retrieving biological data (Cohen, 2004). Sequence alignment becomes an essential tool of bioinformatics and it is vital in various tasks such as genomic annotation, protein secondary and tertiary structure prediction, phylogenetic

DOI: 10.4018/978-1-7998-1204-3.ch031

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tree construction, modeling binding sites, homology searches, gene regulation networks and functional geneomics (Das, Abraham, & Konar, 2008 ; Durbin, Eddy, & Krogh, 1998). From biological point of view all organisms have a common ancestors and so the similarity between DNA or protein sequences exist. The function of newly known sequences with a known sequence can be known with measuring the similarity (Alberts et al., 2007; Arthur, 2002 ; Zvelebil & Baum, 2008).

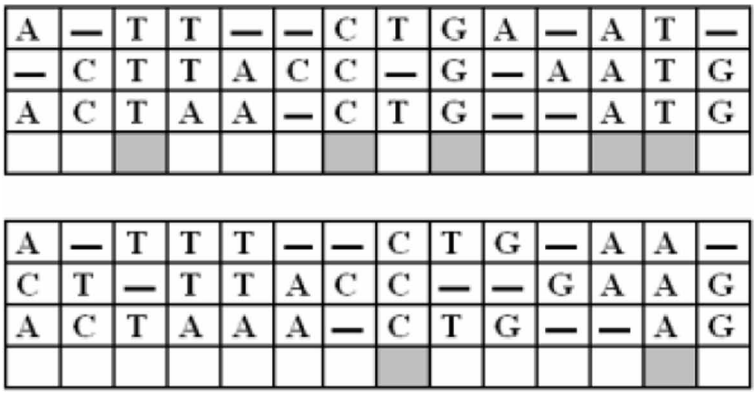
Sequence alignment arranges DNA, RNA and protein sequences to locate conserved blocks or region of similarity. It lining up the nucleotides (A,C,G and T) in DNA or amino acids (20 different amino acids) in protein sequences to achieve the maximum possible level of similarity (Song.J, Liu, Song.Y, & Qu, 2007). The function similarity between sequences is predicted corresponding to the regions of similarity. This arrangement needs insertion of gaps in positions that maximize the alignment score and nucleotides/residues matching.

Finding sequence alignment experimentally is sensitive to less accuracy due to experimental errors with much time consuming and cost. Hence, many efforts in the last years to develop software tools that propose efficient model for accurate alignment. Aligning two sequences is called pairwise sequence alignment. While aligning more than two sequences is called multiple sequence alignment (MSA) as shown in Figure 1 (Sievers & Higgins, 2014). MSAs computation is almost computationally expensive and it classified as NP-complete problem. This chapter focus on the MSA techniques.

The MSA's methods are divided into four approaches: Exact, Progressive, Consistency based and iterative approach (Notredame, 2002). In the exact method (DP) was used for pairwise global alignment by computing the alignment over the entire length of the sequences, (Needleman & Wunsch, 1970). In DP a matrix is created and filled with the partial alignment scores of the two sequences. DP tries to find the shortest path with maximum alignment cost between the start and end of the sequences. The main limitations of DP approach are time and space complexities especially for number of sequences more than 2 sequences (Lipman, Altschul, & Kececioglu, 1989; Carrillo & Lipman, 1988)

Progressive approach solve the problems of the exact method by decreasing the time and space complexties (Taylor, 1988 ; Feng & Doolittle, 1987) The idea of using progressive technique is aligning the most related sequences and then incrementally adding the more distant one by one. The common MSA techniques that based on the progressive approach are CLUSTALW (Thompson, Higgins, &Gibson, 1994), MUSCLE (Edgar, 2004), CLUSTAL OMEGA (Sievers & Higgins, 2014) and Multi-Align

Figure 1. Example of Aligning 3 DNA sequences (MSA)
Öztürk & Aslan, 2016.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(Devereux, Haeberli, & Smithies, 1984). The limitations of progressive approach are that final results depend on the initial pairwise sequence alignment and the alignment scoring scheme used. Besides, the time complexity depends mainly on the number of aligned sequences. Iterative and consistency based approaches outperform the progressive alignment in the point of accuracy. Iterative approach based on dividing the alignment into sub-alignment and re-alignment the sub-alignment. The main techniques based on iterative approach are MAFFT (Katoh, Misawa, Kuma, & Miyata, 2002), DALIGN (Morgenstern, Dress, & Werner, 1996), T-COFEE (Notredame, Higgins, & Heringa, 2000) and MUSCLE (Edgar, 2004).

In the consistency based approach the alignment not depend only on the scoring function used but on the behavior of other alignments also. For example, if three sequences A, B and C are to be aligned, if residue A_i is aligned with residue B_i . If residue Ai aligned with residue C_i then this implies residue B_i can be aligned with residue C_i. The limitation of iterative and consistency based approaches is the time complexity. Where the consistency approach has time complexity of $O(k^3n^2)$ where k is number of aligned sequences and n is the maximum length of aligned sequences. The most common techniques based on consistency approach are PCMA (Pei, Sadreyev, & Grishin, 2003), ProbCons (Do, Mahabhashyam, Brudno, & Batzoglou, 2005) and MUMMALS (Pei & Grishin, 2006). The accuracy of the alignment is reflected by computing the scoring function of MSA. This scoring function measure the performance of the technique and two scoring schemes are used:

- Column Score (CS) which is used to increase the number of matched column and is computed using Equation 1.

$$CS = \sum_{i=1}^{L} M_t(1 + \frac{M_t}{n}) \tag{1}$$

where M_t is the number of matched residues, n is the number of aligned sequences and L is the number of columns in the alignment.

- Sum of Pair (SOP) is the second scoring scheme used for increasing the similarity between two aligned sequences and is computed using Equation 2.

$$SOP = \sum_{i=1}^{L} Score(S(i)_1, S(i)_2) \tag{2}$$

where S_1 and S_2 are the aligned sequences.

For proteins, there are matrices give weight for the bases comparisons. The most common ones are BLOcks of amino acid SUbstitution Matrix (BLOSUM) (Henikoff, S., & Henikoff, J. G., 1992) and Point Accepted Mutation (PAM) (Altschul et al., 1990). There are penalty for contagious gaps instead of separated gaps, so the general gap penalty is used. g_0 and g_e are the gap open and gap extension penalties (Gotoh, 1982).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

Meta-heuristics is a branch of computational intelligence which use heuristics techniques to explore the feasible search space in efficient way. Two main categories of meta-heuristics: single based and population based methods. Single based methods use only one solution and improve it until reach its optimum. The common techniques for single based solutions are Tabu Search (TS) (Glover, 1986), Simulated Annealing (SA) (Kirkpatrick, Gelatt, & Vecchi, 1983), Variable Neighborhood Search (VNS) (Mladenović, 1997 ; Mladenovic, 1995) and Iterated Local Search (ILS) (Stützle, 1999). Population based techniques use many solutions to find the optimum one. The most common ones are Particle Swarm Optimization (PSO) (Kennedy, 2011), Ant Colony Optimization (ACO) (Dorigo, 1992), Genetic Algorithms (GA) (Aryanezhad & Hemati, 2008), Firefly algorithm (Yang, 2010a), Bat-inspired algorithm (Yang, 2010b), Bacterial for aging Optimization (Carillo and Lipman, 1988) and Scatter Search (SS) (Glover, 1998). Two terms related the optimization techniques diversification and intensification. Diversification represents the exploration of the feasible search space. While, intensification represents focusing in certain region of feasible search space. Single based techniques have the property of intensification while population based solution have the property of diversification.

MAIN FOCUS OF THE CHAPTER

MSA problem was solved using simulated annealing like (Kim, Pramanik, & Chung, 1994 ; Simons, Kooperberg, Huang, &Baker, 1997 ; Lukashin, Engelbrecht, &Brunak, 1992). The main problem of simulated annealing is the slow of execution but it well as alignment improver. Genetic algorithm also was used for optimizing MSA in many work like genetic algorithm by optimizing the alignment scoring function (Bucak & Uslan, 2011 ; Botta & Negro, 2009], Progressive Alignment Method using Genetic Algorithm (Naznin, Sarker, & Essam, 2012), Multiple Sequence Alignment using Genetic Algorithm (MSA-GA) (Gupta, Agrwal, & Soni, 2013), Rubber Band Technique using Genetic Algorithm (RBT-GA) (Taheri & Zomaya, 2009) and Vertical Decomposition with Genetic Algorithm (VDGA) (Naznin,, Sarker, & Essam, 2011). Another development using genetic algorithm was using the multi-group and migration technique(Luo, Zhang, & Liang, 2011). Genetic algorithm alignment works well for small number of sequences. With increasing the number of aligned sequences the exponential growth of time is observed. Particle swarm optimization was used for optimizing MSA and proved its superiority in speed of convergence than simulated annealing and genetic algorithm. Also, its ability for aligning sequences more than genetic algorithm. The reason of PSO' success are it has a few parameters to adjust and it needs low memory allocation. A lot of research work proved these properties like (Xu & Chen, 2009; Lalwani, Kumar, & Gupta, 2013; Yadav & Banka, 2016 ; Moustafa, Elhosseini, Taha, & Salem, 2016). Ant Colony Optimization (ACO) and Artificial Bee Colony (ABC) were used for optimizing MSA (Lee & King, 2014 ; Rubio-Largo, Vega-Rodríguez, & González-Álvarez, 2016 ; Öztürk & Aslan, 2016).

BFO is one of the recent nature inspired optimization techniques that mimics bacteria (swarm of E.coli bacteria) for searching for nutrients (Passino, 2012). During bacteria foraging locomotion is achieved by a group of tensile flagella. Flagella lead to tumble or swim operation for the E.coli bacteria. These two operation are performed during the bacteria foraging. Each flagellum pulling on the cell during the bacteria rotating the flagella in the clockwise direction. This action lead to reduce the number of bacteria tumbling. However, in the harm places the tumbling process is increased to find a nutrient gradient.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rotating the flagella in the counter-clockwise direction lead to swimming of the bacteria at a fast rate. Clockwise and counter-clockwise movements take place to find a nutrient solution. When the bacteria get enough food they are increased. From this phenomenon BFO was inspired. A review for optimizing some applications using BFO exists (Das, Biswas, Dasgupta, & Abraham, 2009). In the following FTLPSO and MO-BFO techniques are presented.

FTLPSO

Fragmented protein sequence alignment using two-layer particle swarm optimization (FTLPSO) (Moustafa, Elhosseini, Taha, & Salem, 2016) is presenting efficient result for optimizing MSA using particle swarm optimization tecniques. FTLPSO based mainly on fragmenting the long sequences to prevent PSO from trapping in local minima. In addition, The two layer (main layer and secondary layer) also increase the local minima immunity. The main layer uses the column score objective function while the secondary uses the sum of pair objective function. Each particle in the main layer has many particles in the secondary layer that is as shown in Figure 2. Each particle in the secondary layer point to the gap positions in a fragment. Two vectors of each particle the first one contains the positions of gaps in each fragment. The second vector has a dimension as the number of aligned sequences. Each element of the vector has the value of the number of consecutive gaps in the first vector to be inserted in the corresponding aligned sequence fragment. For example, for a particle *m* assume the first vector contain the following first vector: *m=[3,8,10,1,6,9,10]* and the second vector is: μ *= [1,2,4].* As shown in Figure 3 in the right three fragments of sequences after insertion the gaps represented by one particle. In the secondary layer, the populations in the same group each iteration determine the global best particle. At the end of each iteration this best particle is mutated to reduce the risk of trapping in the local minima. After finishing the iterations of secondary layer the global best particle of each group become the head of this group in the main layer. So, the main layer contains the global best particles of all groups in the secondary layer. The best particle of the global best particles is determined to be the overall global best and the iteration of secondary layer is repeated guiding this new global best particle.

Comparison between FTLPSO and standard PSO was performed using Balibase dataset of different lengths. For short lengths PSO and FTLPSO success to find the same score but for long sequences FTLPSO outperform standard PSO as shown in Figure 4.

In all, this tool is a reliable for finding alignment of MSA. Comparison between it and the common tools for MSA alignment (CLUSTAL OMEGA, TCOFFEE and KALIGN) shows verification of the result as shown in Table 1. The reasons of success of this tool are: first, the fragmentation and mutation reduce the risk of trapping in local minima. Second, Failing of one particle to don't find the optimum score don't affect on the other particles.

MO-BFO

Most of the available technique for optimizing the MSA are used on of the two common objective functions CS and SOP or hybrid between it. So, the objective of is to maximize the objective function which is called unary objective optimization. Another category of optimization for multi objectives as this technique presented. The objectives that were used to optimize MSA are:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The two layer particle swarm optimization techniques
Moustafa, Elhosseini, Taha, & Salem, 2016.

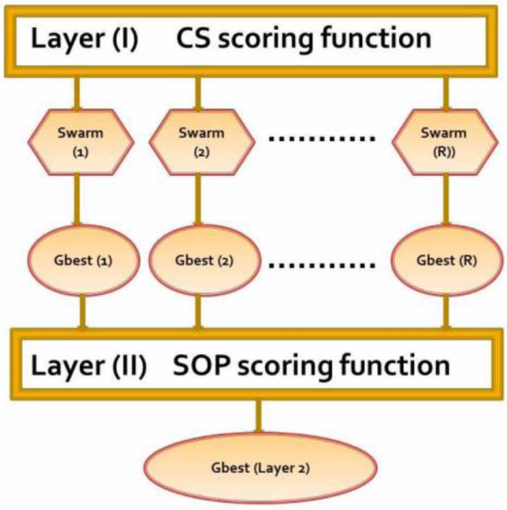
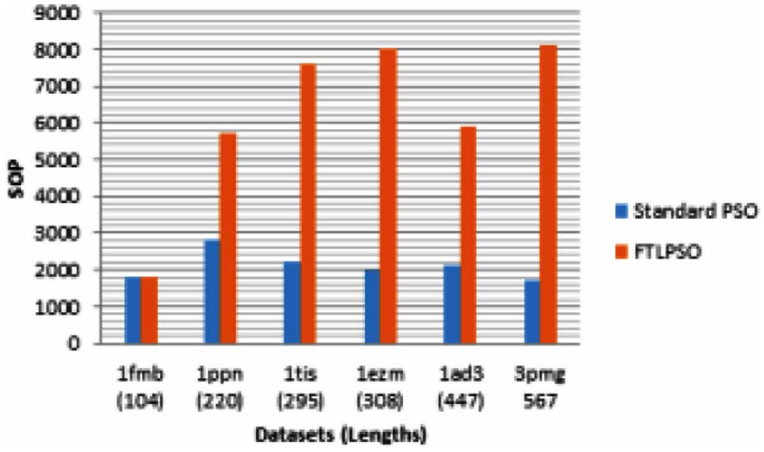


Figure 3. Gap insertion in three fragments of sequences
Moustafa, Elhosseini, Taha, & Salem, 2016.

Aligned Seq.		1	2	3	4	5	6	7	8	9	10	11
ACTCCTGTAC	→	A	C	—	T	C	C	T	G	T	A	C
ACCTCGTTC		A	C	C	T	C	G	T	—	T	—	C
CCTCTGC		—	C	C	T	C	—	T	G	—	—	C

$m = [3,8,10,1,6,9,10]$ and $\mu = [1,2,4]$

Figure 4. Comparison between Standard PSO and FTLPSO using Balibase datasets
Moustafa, Elhosseini, Taha, & Salem, 2016.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Comparison between FTLPSO and three common tools for MSA

Datasets (Length)	FTLPSO	CLUSTAL OMEGA	TCOFFEE	KALIGN
1fmb (104)	1495	1465	1470	1465
1ppn (220)	5256	5156	5243	5230
1tis (295)	7224	7225	7156	7126
1ezm (308)	8211	8110	8141	8233
1ad3 (447)	5632	5487	5681	5681

Moustafa, Elhosseini, Taha, & Salem, 2016.

1. **Maximization of Similarity:** The similarity was estimated using the following formula:

$$Similarity = \frac{dv(m)}{l}$$

$$dv(m) = \max_{n} \left\{ f(n,m) \right\}, m=1,2,3,4,\ldots,l$$

where f(n,m) is the score value of amino acid or nucleotide n on the column m. l is the sequence alignment length.

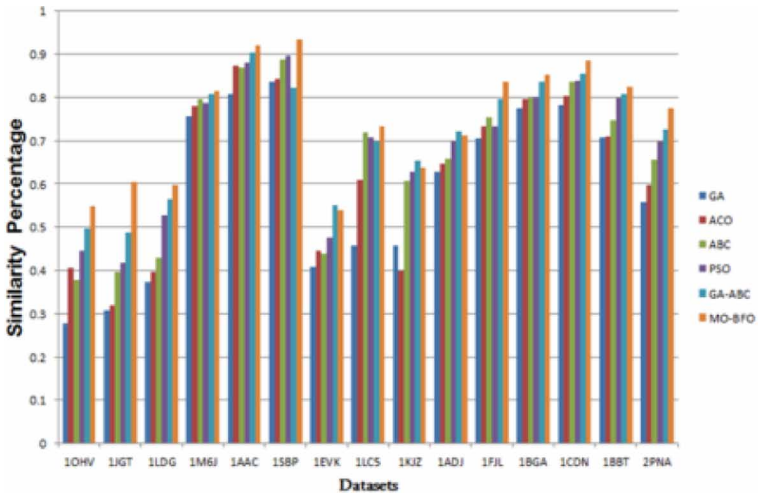
2. **Minimization Gap Penalty:** This objective mean reducing the insertion of gaps as possible that is due to this leads to increasing the alignment score. Column of aligned gap is not allowed. Affine and variable length gap scheme were used in this technique.
3. **Maximize Non-Gap Percentage:** Non gap percentage is defined as the total number of amino acids in the sequence with respect to the number of gaps in the sequences.
4. **Maximize Conserved Blocks:** Regions of the aligned sequences are conserved or identical are called conserved blocks. These region are suitable for phylogenetic analysis or determine common function and structure.

The proposed technique was tested using Balibase datasets version 3.0 and compared with five optimized MSA techniques Genetic Algorithm, ABC, ACO, PSO and hybrid between genetic algorithm and ABC. Figure 5 shows the similarity percentage of various datasets with different lengths. As shown MO-BFO outperform other techniques in almost all datasets. The technique depend on the nature of the sequences if the sequences in the datasets have high difference in length this leads to low similarity and accuracy as shown for datasets *1EKV*, *1KJZ* and *1ADJ*.

Figure 6 shows the result of Non-Gap percentage for the same datasets. MO-BFO has the greatest percentage overall. Gap penalty for all datasets are the minimum for MO-BFO as shown in Figure 7. The performance of MO-BFO was measured by estimating SOP and TCS alignment scores. Table 2 shows the result of comparison of SOP and TCS alignment scores with GA, ACO, ABC, PSO and GA-ABC. The comparisons with online tools such as MAFFT, CLUSTAL OMEGA, K-ALIGN and MUSCLE as shown in Table 3. As shown in Tables 2 & 3 MO-BFO outperform the other methods in term of performance which verify its efficiency. Maximization of conserved blocks is the last objective

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Similarity Percentage of MO-BFO compared with other optimization approaches for MSA was tested on Balibase datasets



of MO-BFO. The technique succeeded to maximize the number of conserved blocks with near optimal scores for other objectives (Similarity, Gap penalty and Non Gap percentage). For example, alignment the dataset 1SBP as shown in Figure 8 shows the conserved patterns in the alignment with non-gap percentage 87% and similarity 93%.

CONCLUSION

This chapter review the recent research for developing MSA using Meta-heuristics or stochastic techniques. Stochastic techniques success in a lot of research to compute the accurate alignment near to that of DP technique. Also, chapter focus deeply on two techniques for MSA are FTLPSO and MO-BFO. FTLPSO use two layer PSO for MSA that decrease the probability of trapping in local minima. MO-BFO use the bacterial foraging technique for optimizing multi objectives of MSA. This chapter conclude the methodology of the two techniques and their performance is compared with other stochastic MSA METHODS and online methods.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Non-gap percentage of MO-BFO compared with other optimization techniques was tested on Balibase datasets

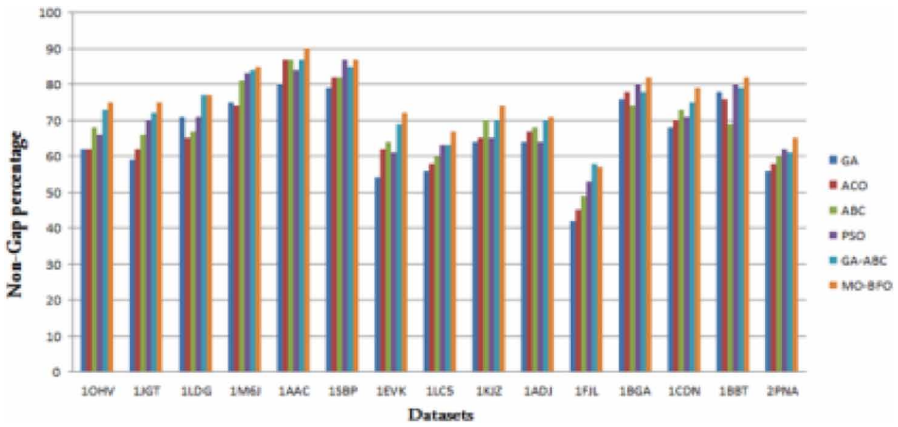


Figure 7. Gap penalty of MO-BFO compared with other optimization techniques was tested on Balibase datasets

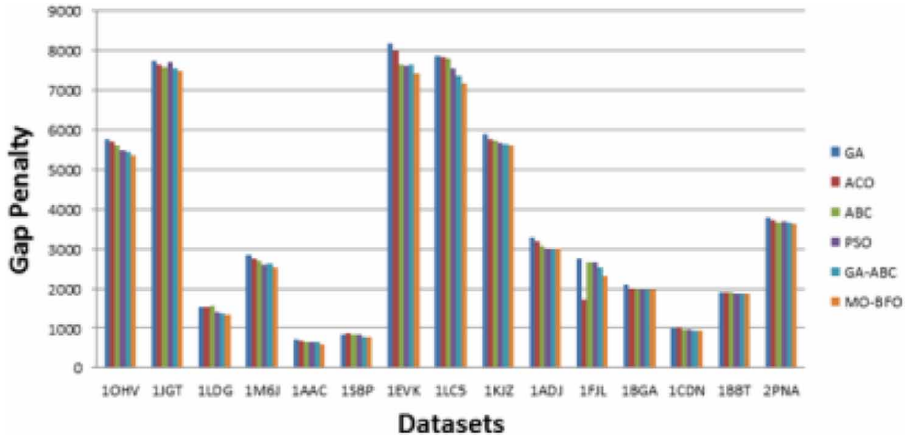


Figure 8. Conserved blocks for alignment of 1SBP dataset using MO-BFO

Seq #	Aligned Sequences														
1	-	K	D	I	Q	L	L	N	V	K	Q	E	T	G	D
2	-	-	E	L	K	V	V	-	-	T	A	E	T	G	H
3	N	N	T	L	Y	F	Y	-	-	F	T	E	T	G	I
1	V	V	I	D	Q	S	H	-	-	-	I	D	K	N	W
2	V	V	I	S	S	-	-	-	-	-	I	D	K	-	N
3	V	V	I	-	T	Y	E	Q	K	-	I	D	K	S	L
1	P	N	P	K	S	S	G	G	D	V	L	I	K	K	L
2	N	Q	G	K	V	L	A	G	N	G	W	R	K	K	L
3	Q	T	A	S	G	A	A	D	L	G	F	V	K	K	L

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. SOP and TCS alignment scores of MO-BFO compared with other optimization techniques

Dataset	Performance Measure	GA	ACO	ABC	PSO	GA- ABC	MO-BFO
1OHV	SOP CS	69 51	70 61	75 60	76 65	76 68	86 73
1JGT	SOP CS	50 40	63 50	61 49	65 47	66 49	69 53
1LDG	SOP CS	74 61	79 64	79 68	76 65	82 70	85 72
1M6J	SOP CS	60 49	65 51	66 54	76 51	78 53	80 60
1AAC	SOP CS	83 72	83 72	87 78	88 76	87 78	92 81
1SBP	SOP CS	86 71	86 78	90 78	91 81	92 83	94 85
1EVK	SOP CS	69 58	74 60	70 62	75 64	79 64	82 69
1LC5	SOP CS	67 54	65 53	68 54	69 59	68 57	71 59
1KJZ	SOP CS	68 55	76 59	76 62	74 61	83 64	87 78
1ADJ	SOP CS	73 54	69 54	79 61	80 65	82 71	85 69
1FJL	SOP CS	70 59	73 60	74 62	78 58	81 68	80 71
1BGA	SOP CS	73 65	76 67	79 69	76 65	83 69	86 72
1CDN	SOP CS	86 76	85 69	87 73	89 78	90 82	93 82
1BBT	SOP CS	74 63	78 64	77 67	79 67	82 69	84 70
2PNA	SOP CS	60 49	59 48	63 52	65 51	64 54	67 52

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. SOP and CS alignment scores of MO-BFO compared with online MSA techniques

Dataset	Performance Measure	CLUSTAL OMEGA	MAFFT	MUSCLE	KALIGN	MO-BFO
1OHV	SOP CS	65 58	60 49	69 55	66 51	86 73
1JGT	SOP CS	49 36	50 36	55 49	61 40	69 53
1LDG	SOP CS	77 65	78 67	77 60	70 59	85 72
1M6J	SOP CS	79 56	77 57	80 66	75 53	80 60
1AAC	SOP CS	86 73	85 72	78 68	73 65	92 81
1SBP	SOP CS	90 82	87 73	91 83	84 69	94 85
1EVK	SOP CS	79 66	60 56	77 69	51 40	82 69
1LC5	SOP CS	59 49	64 53	69 52	64 53	71 59
1KJZ	SOP CS	68 54	66 54	64 52	65 50	87 78
1ADJ	SOP CS	79 53	76 57	78 58	74 60	85 69
1FJL	SOP CS	73 43	75 52	79 56	74 51	80 71
1BGA	SOP CS	84 68	78 69	79 64	75 63	86 72
1CDN	SOP CS	91 84	86 79	90 76	85 75	93 82
1BBT	SOP CS	84 58	74 60	79 63	76 64	84 70
2PNA	SOP CS	61 43	57 44	61 48	53 42	67 52

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *General principles of cell communication*. Academic Press.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2007). *Molecular Biology of the Cell*. London: Garland Science.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. doi:10.1016/S0022-2836(05)80360-2 PMID:2231712

Gotoh, O. (1982, December). An improved algorithm for matching biological sequences. *Journal of Molecular Biology*, *162*(3), 705–708. doi:10.1016/0022-2836(82)90398-9 PMID:7166760

Arthur, L. M. (2002). Introduction to Bioinformatics. Oxford University Press.

Aryanezhad, M. B., & Hemati, M. (2008). A new genetic algorithm for solving nonconvex nonlinear programming problems. *Applied Mathematics and Computation*, *199*(1), 186–194. doi:10.1016/j.amc.2007.09.047

Botta, M., & Negro, G. (2009, October). Multiple sequence alignment with genetic algorithms. In *International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics* (pp. 206-214). Springer Berlin Heidelberg.

Bucak, I. Ö., & Uslan, V. (2011). Sequence alignment from the perspective of stochastic optimization: A survey. *Turkish Journal of Electrical Engineering & Computer Sciences*, *19*(1), 157–173.

Carrillo, H., & Lipman, D. (1988). The multiple sequence alignment problem in biology. *SIAM Journal on Applied Mathematics*, *48*(5), 1073–1082. doi:10.1137/0148063

Chen, H., Zhu, Y., & Hu, K. (2009). Cooperative bacterial for aging optimization. *Discrete Dynamics in Nature and Society*.

Cohen, J. (2004). Bioinformatics—an introduction for computer scientists. *ACM Computing Surveys*, *36*(2), 122–158. doi:10.1145/1031120.1031122

Das, S., Abraham, A., & Konar, A. (2008). Swarm intelligence algorithms in bioinformatics. In *Computational Intelligence in Bioinformatics* (pp. 113–147). Springer Berlin Heidelberg. doi:10.1007/978-3-540-76803-6_4

Das, S., Biswas, A., Dasgupta, S., & Abraham, A. (2009). Bacterial foraging optimization algorithm: theoretical foundations, analysis, and applications. In *Foundations of Computational Intelligence* (Vol. 3, pp. 23–55). Springer Berlin Heidelberg. doi:10.1007/978-3-642-01085-9_2

Devereux, J., Haeberli, P., & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research*, *12*(1), 387–395. doi:10.1093/nar/12.1Part1.387 PMID:6546423

Do, C. B., Mahabhashyam, M. S., Brudno, M., & Batzoglou, S. (2005). ProbCons: Probabilistic consistency-based multiple sequence alignment. *Genome Research*, *15*(2), 330–340. doi:10.1101/gr.2821705 PMID:15687296

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Dorigo, M. (1992). *Optimization, learning and natural algorithms* (Ph. D. Thesis). Politecnico di Milano, Italy.

Durbin, R., Eddy, S. R., Krogh, A., & Mitchison, G. (1998). *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge University Press. doi:10.1017/CBO9780511790492

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*(1), 113. doi:10.1186/1471-2105-5-113 PMID:15318951

Feng, D. F., & Doolittle, R. F. (1987). Progressive sequence alignment as a prerequisitetto correct phylogenetic trees. *Journal of Molecular Evolution*, *25*(4), 351–360. doi:10.1007/BF02603120 PMID:3118049

Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, *13*(5), 533–549. doi:10.1016/0305-0548(86)90048-1

Glover, F. (1998). A template for scatter search and path relinking. *Lecture Notes in Computer Science*, *1363*, 13–54. doi:10.1007/BFb0026589

Gupta, R., Agarwal, P., & Soni, A. K. (2013). MSA-GA: Multiple Sequence Alignment Tool Based On Genetic Approach. *Int. J. Soft Comput. Software Eng.*, *3*(8), 1–11.

Henikoff, S., & Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(22), 10915–10919. doi:10.1073/pnas.89.22.10915 PMID:1438297

Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. doi:10.1093/nar/gkf436 PMID:12136088

Kim, J., Pramanik, S., & Chung, M. J. (1994). Multiple sequence alignment using simulated annealing. Computer applications in the biosciences. *CABIOS*, *10*(4), 419–426. PMID:7804875

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*(4598), 671-680.

Lalwani, S., Kumar, R., & Gupta, N. (2013). A review on particle swarm optimization variants and their applications to multiple sequence alignment. *Journal of Applied Mathematics and Bioinformatics*, *3*(2), 87.

Lee, A., & King, L. (2014). *ACO Implementation for Sequence Alignment with Genetic Algorithms*. arXiv preprint arXiv:1406.0930.

Lipman, D. J., Altschul, S. F., & Kececioglu, J. D. (1989). A tool for multiple sequence alignment. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(12), 4412–4415. doi:10.1073/pnas.86.12.4412 PMID:2734293

Lukashin, A. V., Engelbrecht, J., & Brunak, S. (1992). Multiple alignment using simulated annealing: Branch point definition in human mRNA splicing. *Nucleic Acids Research*, *20*(10), 2511–2516. doi:10.1093/nar/20.10.2511 PMID:1598209

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Luo, J., Zhang, L., & Liang, C. (2011, September). A multigroup parallel genetic algorithm for multiple sequence alignment. In *International Conference on Artificial Intelligence and Computational Intelligence* (pp. 308-316). Springer Berlin Heidelberg. 10.1007/978-3-642-23881-9_41

Mladenovic, N. (1995, May). A variable neighborhood algorithm-a new metaheuristic for combinatorial optimization. In *Papers presented at Optimization Days* (vol. 12).

Mladenović, N., & Hansen, P. (1997). Variable neighborhood search. *Computers & Operations Research*, *24*(11), 1097–1100. doi:10.1016/S0305-0548(97)00031-2

Morgenstern, B., Dress, A., & Werner, T. (1996). Multiple DNA and protein sequence alignment based on segment-to-segment comparison. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(22), 12098–12103. doi:10.1073/pnas.93.22.12098 PMID:8901539

Moustafa, N., Elhosseini, M., Taha, T. H., & Salem, M. (2016). Fragmented protein sequence alignment using two-layer particle swarm optimization (FTLPSO). *Journal of King Saud University-Science*.

Naznin, F., Sarker, R., & Essam, D. (2011). Vertical decomposition with genetic algorithm for multiple sequence alignment. *BMC Bioinformatics*, *12*(1), 353. doi:10.1186/1471-2105-12-353 PMID:21867510

Naznin, F., Sarker, R., & Essam, D. (2012). Progressive alignment method using genetic algorithm for multiple sequence alignment. *IEEE Transactions on Evolutionary Computation*, *16*(5), 615–631. doi:10.1109/TEVC.2011.2162849

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, *48*(3), 443–453. doi:10.1016/0022-2836(70)90057-4 PMID:5420325

Notredame, C. (2002). Recent progress in multiple sequence alignment: A survey. *Pharmacogenomics*, *3*(1), 131–144. doi:10.1517/14622416.3.1.131 PMID:11966409

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, *302*(1), 205–217. doi:10.1006/jmbi.2000.4042 PMID:10964570

Öztürk, C., & Aslan, S. (2016). A new artificial bee colony algorithm to solve the multiple sequence alignment problem. *International Journal of Data Mining and Bioinformatics*, *14*(4), 332–353. doi:10.1504/IJDMB.2016.075823

Particle Swarm Optimization. (n.d.). In *Encyclopedia of machine learning* (pp. 760-766). Springer US.

Passino, K. M. (2012). Bacterial foraging optimization. *Innovations and Developments of Swarm Intelligence Applications*, 219-233.

Pei, J., & Grishin, N. V. (2006). MUMMALS: Multiple sequence alignment improved by using hidden Markov models with local structural information. *Nucleic Acids Research*, *34*(16), 4364–4374. doi:10.1093/nar/gkl514 PMID:16936316

Pei, J., Sadreyev, R., & Grishin, N. V. (2003). PCMA: Fast and accurate multiple sequence alignment based on profile consistency. *Bioinformatics (Oxford, England)*, *19*(3), 427–428. doi:10.1093/bioinformatics/btg008 PMID:12584134

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rubio-Largo, Á., Vega-Rodríguez, M. A., & González-Álvarez, D. L. (2016). Hybrid multiobjective artificial bee colony for multiple sequence alignment. *Applied Soft Computing*, *41*, 157–168. doi:10.1016/j.asoc.2015.12.034

Sievers, F., & Higgins, D. G. (2014). Clustal Omega, accurate alignment of very large numbers of sequences. *Multiple Sequence Alignment Methods*, 105-116.

Simons, K. T., Kooperberg, C., Huang, E., & Baker, D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *Journal of Molecular Biology*, *268*(1), 209–225. doi:10.1006/jmbi.1997.0959 PMID:9149153

Song, J., Liu, C., Song, Y., & Qu, J. (2007, December). Alignment of multiple proteins with an ensemble of Hidden Markov Models. In *Machine Learning and Applications, 2007. ICMLA 2007. Sixth International Conference on* (pp. 594-599). IEEE. 10.1109/ICMLA.2007.90

Stützle, T. G. (1999). *Local search algorithms for combinatorial problems: analysis, improvements, and new applications* (Vol. 220). Sankt Augustin: Infix.

Taheri, J., & Zomaya, A. Y. (2009). RBT-GA: A novel metaheuristic for solving the multiple sequence alignment problem. *BMC Genomics*, *10*(Suppl 1), S10. doi:10.1186/1471-2164-10-S1-S10 PMID:19594869

Taylor, W. R. (1988). A flexible method to align large numbers of biological sequences. *Journal of Molecular Evolution*, *28*(1-2), 161–169. doi:10.1007/BF02143508 PMID:3148736

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. doi:10.1093/nar/22.22.4673 PMID:7984417

Xu, F., & Chen, Y. (2009, September). A method for multiple sequence alignment based on particle swarm optimization. In *International Conference on Intelligent Computing* (pp. 965-973). Springer Berlin Heidelberg. 10.1007/978-3-642-04020-7_104

Yadav, R. K., & Banka, H. (2016). A PSO with Improved Initialization Operator for Solving Multiple Sequence Alignment Problems. In *Proceedings of the 4th International Conference on Frontiers in Intelligent Computing: Theory and Applications (FICTA) 2015* (pp. 285-294). Springer India. 10.1007/978-81-322-2695-6_25

Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010) (pp. 65-74). Springer Berlin Heidelberg. doi:10.1007/978-3-642-12538-6_6

Yang, X. S. (2010). Firefly algorithm, stochastic test functions and design optimisation. *International Journal of Bio-inspired Computation*, *2*(2), 78–84. doi:10.1504/IJBIC.2010.032124

Zvelebil, M., & Baum, J. O. (2008). *Understanding bioinformatics* (Doctoral dissertation). Garland Science, Taylor & Francis Group, LLC.

This research was previously published in the Handbook of Research on Machine Learning Innovations and Trends edited by Aboul Ella Hassanien and Tarek Gaber; pages 409-423; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 32
GEView (Gene Expression View) Tool for Intuitive and High Accessible Visualization of Expression Data for Non-Programmer Biologists

Libi Hertzberg
Tel Aviv University, Israel & Weizmann Institute of Science, Israel

Assif Yitzhaky
Weizmann Institute of Science, Israel

Metsada Pasmanik-Chor
Tel Aviv University, Israel

ABSTRACT

This article describes how the last decade has been characterized by the production of huge amounts of different types of biological data. Following that, a flood of bioinformatics tools have been published. However, many of these tools are commercial, or require computational skills. In addition, not all tools provide intuitive and highly accessible visualization of the results. The authors have developed GEView (Gene Expression View), which is a free, user-friendly tool harboring several existing algorithms and statistical methods for the analysis of high-throughput gene, microRNA or protein expression data. It can be used to perform basic analysis such as quality control, outlier detection, batch correction and differential expression analysis, through a single intuitive graphical user interface. GEView is unique in its simplicity and highly accessible visualization it provides. Together with its basic and intuitive functionality it allows Bio-Medical scientists with no computational skills to independently analyze and visualize high-throughput data produced in their own labs.

DOI: 10.4018/978-1-7998-1204-3.ch032

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Advanced high-throughput technologies have been developed and extensively used in the last decades (microarrays since the mid-1980s (Schena, Shalon, Davis, & Brown, 1995) and Next Generation Sequencing (NGS) (Goodwin, McPherson, & McCombie, 2016) since the mid-2000s). Mass spectrometry technologies were originally invented almost 100 years ago and further developed during the 1990s (Glish & Vachet, 2003). All these technologies and others are being routinely used by Bio-medical (Bio-Med) scientists for production of high-throughput data. However, currently, it is very difficult for Bio-Med researchers with no computational skills to analyze their own data. Various free, online and user-friendly tools have been developed and are routinely used for gene expression and proteomics data analysis. For example, Expander (Ulitsky et al., 2010), Chipster (Kallio et al., 2011), SAM (Tusher, Tibshirani, & Chu, 2001), Limma (Ritchie et al., 2015), DESeq (Love, Huber, & Anders, 2014), Morpheus (https://software.broadinstitute.org/morpheus/), MeV (Howe, Sinha, Schlauch, & Quackenbush, 2011) for gene expression analysis, and MaxQuant (Cox & Mann, 2008) for proteomics data analysis. However, these tools usually perform a large range of tasks, which are sometimes not straightforward and may necessitate some computational skills. In addition, they usually lack a single-gene trivial visualization (such as box plots), which is a basic request of lab researchers in order to study the expression pattern of a specific gene of interest. While the Morpheus tool (https://software.broadinstitute.org/morpheus/) does provide box plots for single genes, it has the limitation of not providing differential expression statistical analysis of more than two groups of samples. The result is that many Bio-Med researchers are reluctant to use such tools and often outsource for the analysis of their data. Outsourcing creates a gap in time and expenditure of financial resources, in addition to the fact that it is performed by a computational expert, who doesn't necessarily have a biological view of results. Thus, there is a need for a more basic and intuitive tool, with functionalities such as quality control measures, simple statistics, differential expression analysis and intuitive visualization of the results, including single-gene visualizations.

GEView is a free, easily downloadable tool suitable for Windows and Linux platforms, written in MATLAB (https://www.mathworks.com). It provides intuitive quality control analysis of the experiment, outlier exclusion and easy identification of differential expression of gene, microRNA or protein data. The great advantage of GEView is its usage simplicity and the output Excel table which provides immediate visualization of the analyzed data on a single-gene basis. It serves all available expression platforms, given that expression values for each gene/microRNA/protein are provided in a tab-delimited file as input, after normalization. GEView is freely available to non-commercial users and can be downloaded from http://www.weizmann.ac.il/complex/compphys/software/geview/. It includes a convenient "help" menu and an example test data (Teuffel et al., 2004).

We believe GEView may be highly useful for Bio-Med researchers who are not computational experts. Such researches, who were previously reluctant to use computational tools and were dependent on external bioinformatics services, have now an opportunity to analyze their own data. In addition, the readily available single-gene graphical representation enables laboratory scientists to simply return to the results table each time they become interested in a new gene or protein and explore its expression pattern in their data. Thus, by improving the access to basic expression analysis, GEView improves the ability of laboratory scientists to produce significant results from their various experiments.

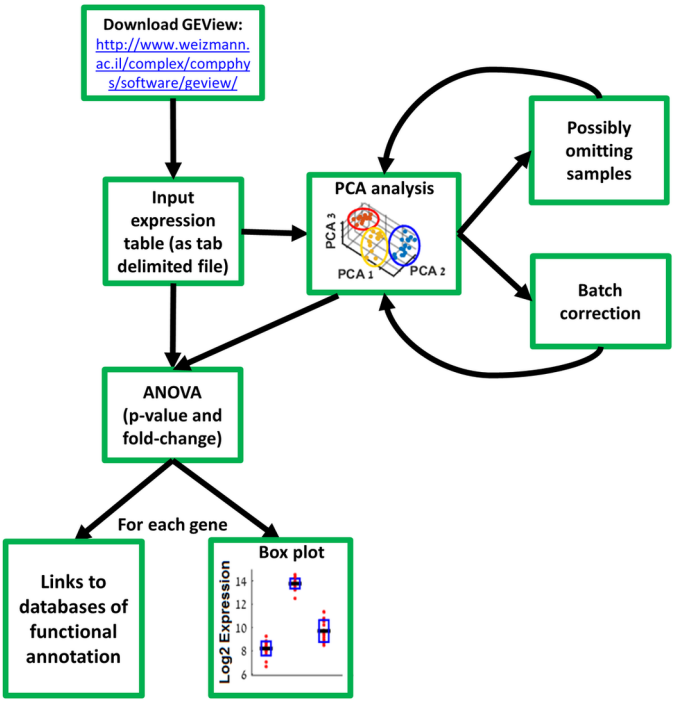
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

IMPLEMENTATION

In this paper we present GEView, a useful tool for various types of expression analysis (genes, microRNAs and proteins). Our aim is to enable the Bio-Med researchers with no computer-science skills, to simply analyze and visualize high-throughput data produced in their wet experiments (see workflow in Figure 1), using an intuitive user interface (see Figure 2).

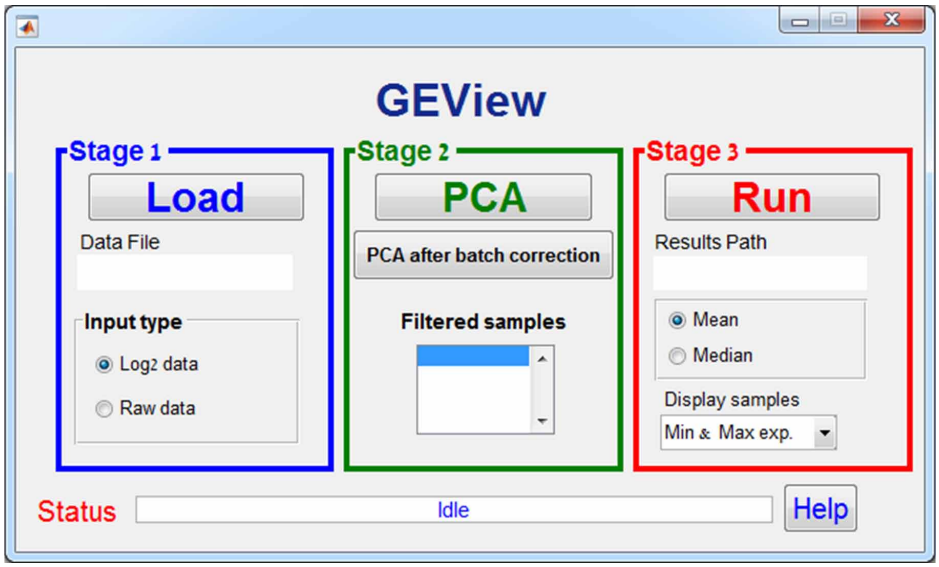
GEView performs a basic quality control step of Principal Component Analysis (PCA) (Jolliffe & Cadima, 2016). The researcher may correlate PCA output with experimental design and sample replicates and examine whether all biological replicates cluster together (otherwise, unknown variables may have affected the data). It may also indicate biological outliers among the samples. In addition, various experimental batches resulting from experimental design and/or other technical constraints can be easily corrected using GEView, implementing the ComBat batch correction algorithm (Johnson, Li, & Rabinovic, 2007; Leek, Johnson, Parker, Jaffe, & Storey, 2012). Moreover, GEView uses Analysis Of Variance (ANOVA) (Bewick, Cheek, & Ball, 2004) (http://www.mathworks.com/help/stats/anova1.html) to identify genes (/microRNAs /proteins) which are differentially expressed between the various experimental conditions. The single output Excel file provides immediate visualization of all genes' differential expression in the various experimental conditions (box plots) together with ANOVA P-values and fold-change values, through a click on each gene-link in the output file. The expression pattern of each single gene can be easily and immediately viewed by clicking on the link in the output file, which opens a pre-computed box plot (for illustration, see an example of a results table in Table 1; the links in the column "Box Plots" open a box plot for each of the genes). In order to complete data examination

Figure 1. GEView flow chart



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. GEView user interface



and analysis, links are also provided for GeneCards (Safran et al., 2010), UCSC (Kent et al., 2002) and NCBI (https://www.ncbi.nlm.nih.gov/gene) databases, for each specific entry (for illustration, see the links in columns 3-5 of Table 1).

RESULTS

Test Data: The "Zurich" Dataset

In order to demonstrate GEView functionality, we use a gene expression dataset of leukemia blood samples that were processed at the department of Oncology, University Children's Hospital, Zurich, Switzerland (Teuffel et al., 2004) and measured by Affymetrix HG-U133A microarrays. We utilize as an example three sample groups that represent three leukemia subtypes, each with different genomic characteristics: E2A-PBX1 (E2A; 5 samples), High hyperdiploid (HD; 4 samples) and TEL-AML1 (TEL; 8 samples). E2A-PBX1 is characterized by a translocation that fuses two genes, E2A and PBX1, High hyperdiploid is characterized by gain of chromosomes (overall 51-67 chromosomes) and TEL-AML1 is characterized by a translocation that fuses two genes, TEL and AML1.

Data Input File and Details for Obtaining It

A tab-delimited text file with normalized expression values for each entry (resulting from microarray, NGS or proteomics) is required as input by GEView. Data can be provided as raw expression data or log 2 transformed, according to the user`s preference. The format is fully described in the "help" document; an example input table is given in Additional File 1 (http://www.weizmann.ac.il/complex/compphys/software/geview/0.4/Additional_File1_Zurich_HD_TEL_E2A_data_labeled.txt).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Various tools are available for free download in order to obtain the normalized expression values. EXPANDER tool (Ulitsky et al., 2010) (http://acgt.cs.tau.ac.il/expander/) or the Expression console tool (for Affymetrix microarray data; https://www.affymetrix.com/support/technical/byproduct.affx?product=expressionconsole) can be used for microarray data, Chipster (Kallio et al., 2011) (http://chipster.csc.fi/) may be used for NGS data, MaxQuant (Cox & Mann, 2008) may be used for proteomics data and R software (https://www.r-project.org/), which requires programing skills, can be used for all types of data. Commercial tools such as Partek Genomics suite (http://www.partek.com/) and MATLAB (https://www.mathworks.com) may also be used.

Load Data

Use the "Load" button (blue; see Figure 2) in order to browse through the input data file in the computer directory. After choosing the data file, the status bar will indicate when its loading is completed.

Principal Component Analysis (PCA) and Batch Effect Correction

In order to evaluate sample quality and homogeneity, PCA (Jolliffe & Cadima, 2016) may be plotted (using the "PCA" green button, see Figure 2), implemented by MATLAB software (http://www.mathworks.com/help/stats/pca.html). The PCA calculation is based on the 1,000 highest variance genes (if there exists less than 1000 genes, then all genes are used). In Figure 3A the PCA of the "Zurich" test data is plotted. It can be seen that one of the TEL (yellow square) samples is clustered within the HD group (red triangles), and not with its own group. Using GEView, the researcher can exclude this sample from the analysis where, for instance, there is evidence that it was mislabeled, by right clicking on it (see Figure 3B). It should be noted that such a decision should be taken carefully, as it has the potential to bias the results.

If various batches are seen by PCA or are known in advance, normalization according to experimental batches may be applied. "PCA after batch correction" button may be used to re-view the new PCA after the batch correction.

Batch Correction Example

When combining datasets from various origins (for example, in case datasets were processed in different dates or laboratories), the samples are often clustered according to the technical origin instead of the experimental conditions. GEView implements the ComBat batch correction algorithm (Johnson et al., 2007; Leek et al., 2012) to enable correction of such technical artifacts. In the following example, we added to our test data that were processed at the department of Oncology, University Children's Hospital Zurich, Switzerland (Teuffel et al., 2004) (labeled "Zurich") an additional dataset, that was processed at St Jude Children's Research Hospital, Memphis, TN (Ross et al., 2003) (labeled "StJude"). These datasets share two of the sample groups, TEL and HD. However, it can be seen in Figure 4A, where the samples are shaped and colored according to data origin (batches: "Zurich" – red triangles, "StJude" – blue circles), that the samples are grouped according to the batches; Zurich to the right and StJude to the left. After applying the batch correction algorithm (using the "PCA after batch correction" button, see Figure 2), it can be seen in Figure 4B that the batch effect disappeared, and the samples are grouped solely according to the experimental conditions (HD and TEL).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. PCA quality control visualization: A) PCA is plotted for the "Zurich" dataset. Each of the experimental conditions is shaped and colored differently: E2A in blue circles, HD in red triangles and TEL in yellow squares. The colored large circles were added in order to emphasize the clustering into 3 biological groups. B) PCA is plotted for the "Zurich" dataset in the same manner as in A), after removing the outlier TEL sample (marked by the text "Removed").

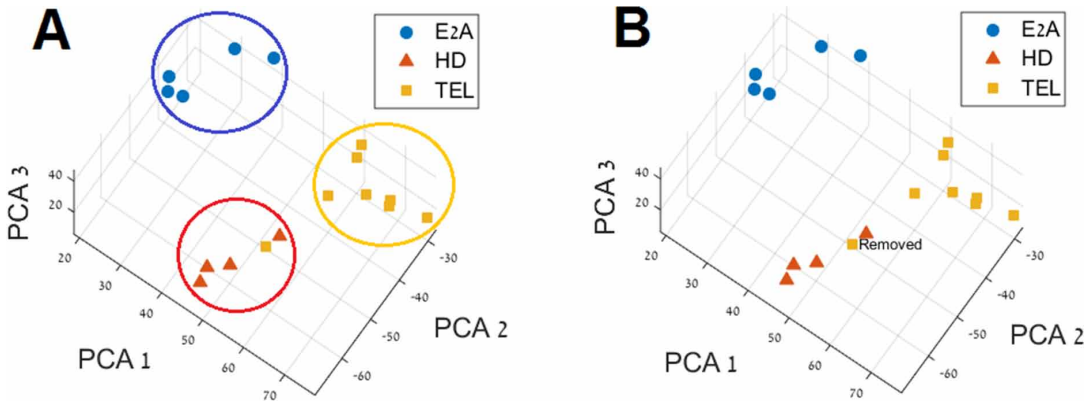
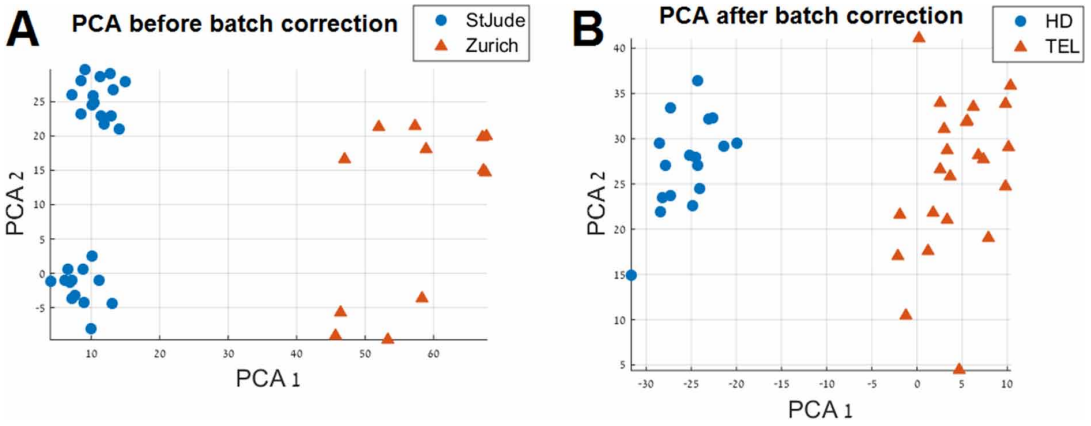


Figure 4. Batch correction example: A) PCA is plotted for the union of HD and TEL samples from both "Zurich" and "StJude" datasets, before the batch correction is applied. The samples are shaped and colored by batch (red triangles for "Zurich" and blue circles for "StJude"). B) PCA is plotted for the union of HD and TEL samples from both "Zurich" and "StJude" datasets after the batch correction was applied. The samples are shaped and colored by experimental conditions (blue circles for HD, red triangles for TEL).



Run Statistical Analysis

By pressing the "RUN" red button (see Figure 2), ANOVA is applied for each probe set (gene/microRNA/protein), testing whether the means of its expression significantly differ between each of the different conditions (see http://www.mathworks.com/help/stats/anova1.html for more details). P-values and Q-values (corrected P-values for multiple comparisons (Benjamini & Hochberg, 1995) are being calculated

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

for each probe set. Running the statistical analysis of a typical dataset containing about 20,000 entries may take up to a few hours, using a computer with basic requirements such as 8GB RAM and Intel Core i7 processor of 3.40 GHz. Running progress can be viewed by the status line at the bottom of the user interface (see Figure 2).

This stage generates a single Excel file which summarizes in each line the results for a specific probe set. Table 1 represents the resulting Excel table for 10 probe sets (see Additional File 2 for the complete ANOVA table for ~ 17,000 probe sets http://www.weizmann.ac.il/complex/compphys/software/geview/0.4/Additional_File2_results_Zurich_HD_TEL_E2A_data_labeled.xlsx).

Exploring the Results

The output single Excel file provides immediate visualization of all genes' differential expression in the various experimental conditions (box plots) together with ANOVA P-values and fold-change values, through a click on each gene-link in the output file (Table 1 contains 10 rows, out of ~17,000, for illustration; the links on the "Box Plots" column can be clicked on). For example, in Figure 5, box plots of two genes are plotted, PBX1 and ELK3.

Each row represents a probe set ID, identified by a gene symbol (second column). Please note that 3 different probe set IDs for PBX1 gene and 2 for ELK3 gene are present. The Box plots column contains a link to a box plot of the mean expression values with standard errors and statistics resulting from ANOVA for each probe set. For each entry, links to GeneCards, UCSC and NCBI genome browser databases are provided. In addition, the same Excel file contains fold change values between each pair of experimental conditions. In case there are more than two experimental conditions, Tukey's test (Tukey JW, n.d.) for multiple comparisons is performed (http://www.mathworks.com/help/stats/multcompare.html), comparing between each pair, and the p-values are given.

As shown in Figure 5, probe sets belonging to the same gene are plotted together. This is done in order to estimate whether the different probe sets that belong to the same gene share a similar expression pattern (if they do, it significantly strengthens the validity of the pattern). It can be seen that PBX1 has 3 probe sets, all of which show a similar and significant expression pattern – high expression in E2A-PBX1 sample group and low in the rest (Figure 5A). E2A-PBX1 leukemia subtype is characterized by a fusion between the transcription factor E2A and PBX1 gene. While normally PBX1 is not expressed in lymphoid cells, the fusion with E2A causes its expression (Aspland, Bendall, & Murre, 2001). Thus, PBX1 is indeed expected to be expressed in E2A-PBX1 sample group and not expressed in the rest. The concordance between all 3 probe sets (Figure 5A) significantly increases the validation of this pattern. Another example of the gene ELK3, can be seen in Figure 5B. It has two probe sets, one showing a significant differential expression ($Q= 0.024$), while the other shows a different, non-significant pattern. Thus, the expression pattern of ELK3 is less validated than that of PBX1. When there is a need to decide on which gene to focus and perform further investigations and wet experiments, this information is crucial. Therefore, the simple grouping of the box plots of different probe sets of the same gene can improve the decision-making process.

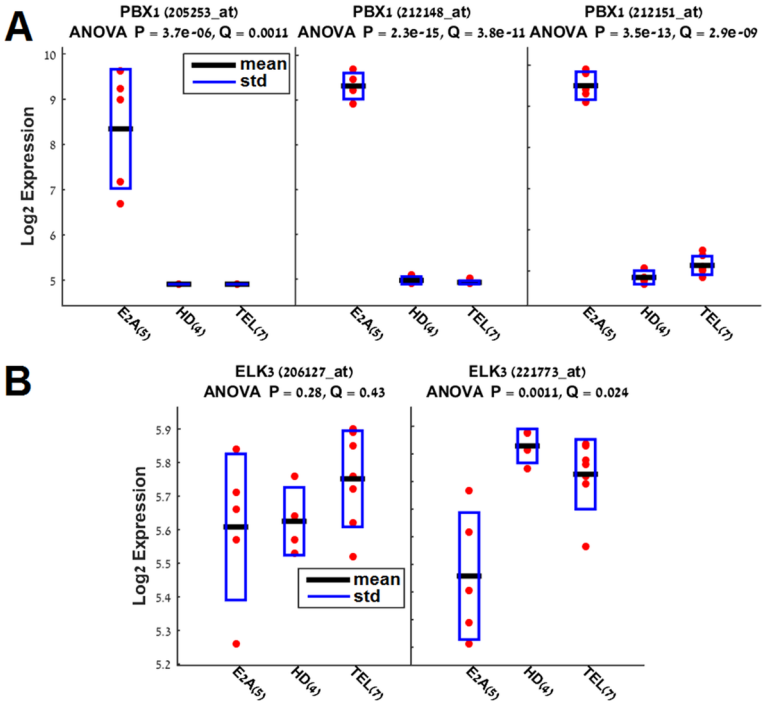
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Output Excel table

Probe Set ID	Gene Symbol	GeneCards Link	UCSC Link	NCBI Link	ANOVA P-Value	Q-Value (Corrected)	Box Plots	E2A/HD Fold-Change	p-Value	E2A/TEL Fold-Change	p-Value	HD/TEL Fold-Change	p-Value
212148_at	PBX1	GeneCards	UCSC	NCBI	2.30E-15	3.80E-11	Figure	20.3	1.90E-09	21	1.90E-09	1.03	0.9
212151_at	PBX1	GeneCards	UCSC	NCBI	3.50E-13	2.90E-09	Figure	25.2	1.90E-09	20.6	1.90E-09	0.817	0.2
200953_s_at	CCND2	GeneCards	UCSC	NCBI	2.10E-07	0.00022	Figure	0.0787	1.70E-07	0.223	1.40E-05	2.83	0.00089
201005_at	CD9	GeneCards	UCSC	NCBI	1.30E-06	0.00055	Figure	0.874	0.92	10.7	6.00E-06	12.3	6.90E-06
205253_at	PBX1	GeneCards	UCSC	NCBI	3.70E-06	0.0011	Figure	10.8	2.60E-05	10.8	6.10E-06	1	1
204849_at	TCFL5	GeneCards	UCSC	NCBI	2.20E-05	0.0029	Figure	0.639	0.34	0.152	2.70E-05	0.238	0.00069
200951_s_at	CCND2	GeneCards	UCSC	NCBI	3.80E-05	0.004	Figure	0.372	3.10E-05	0.744	0.083	2	0.00049
221773_at	ELK3	GeneCards	UCSC	NCBI	0.0011	0.024	Figure	0.195	0.0016	0.279	0.0039	1.43	0.56
200952_s_at	CCND2	GeneCards	UCSC	NCBI	0.13	0.27	Figure	0.789	0.18	0.812	0.18	1.03	0.97
206127_at	ELK3	GeneCards	UCSC	NCBI	0.28	0.43	Figure	0.988	0.99	0.905	0.32	0.916	0.45

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Box plots of PBX1 and ELK3 genes: A) Box plots for the 3 probe set IDs present in the "Zurich" dataset for the gene PBX1 are plotted. In the title of each plot, the gene symbol and probe set ID (in parentheses) are presented. In addition, the ANOVA P and Q values (corrected p-value (Benjamini & Hochberg, 1995)) are provided. The Y-axis represents the Log2 expression value and the X-axis represents the 3 experimental conditions (see E2A, HD and TEL X-labels). The number of samples in each experimental condition is written in parentheses. Each sample expression value is marked by a red dot and the mean and standard deviation in each experimental condition are marked by a black line and a blue box, respectively. B) Box plots for the 2 probe set IDs present in the "Zurich" dataset for the gene ELK3 are plotted in the same manner as in A.



HELP DOCUMENT

A comprehensive Help document can be easily accessed through GEView (press the "Help" button in the right lower corner of the user interface, as plotted in Figure 2; the full Help document is also given as Additional File 3 http://www.weizmann.ac.il/complex/compphys/software/geview/0.4/Additional_File3_help.docx). The Help document contains:

- Installation instructions;
- Input format description for all data types;
- Description of the analysis flow, including a detailed description of the user interface functionality and the running parameters;
- Detailed description of each analysis type provided by GEView, including results of the "Zurich" test dataset.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DISCUSSION

There are many tools that perform gene or protein expression analysis. We offer GEView, a tool that implements basic analyses and intuitive visualization of the results for the non-programmer biologist. The great advantage of GEView is its intuitive, accessible and comprehensive results table, given in an Excel file with links to graphics and databases. Specifically, it provides a single results table, including pre-calculated ANOVA p-values and links to pre-computed figures of the expression pattern of each of the genes/microRNAs/proteins across the experimental conditions. This is a significant improvement to the commonly available results tables, which usually include only p-values and fold-change differences and don't provide visualization of expression patterns of single genes. The intuitive graphical representation, using box plots, enables the laboratory scientist to simply investigate many genes of interest. Gene expression patterns may easily be explored in the same results table, without the need for further external resources, through direct links to GeneCards, UCSC and NCBI databases for each gene (see Table 1).

CONCLUSION

GEView was developed for Bio-Med researchers who are not computational experts. We believe that researchers who were previously dependent on external services for the analysis of their own data, will have a new opportunity to analyze expression results with GEView. Upon input of expression data of various types, GEView provides PCA analysis, batch correction and analysis of differential expression, with immediate graphical view and functional annotation of the results, in a single simple Excel file. While all algorithms used are published, implemented by standard and validated MATLAB functions (in the case of ComBat algorithm, using the original R software code (Leek et al., 2012), wrapping them together in a simple and intuitive tool with immediate access to graphical representation on a single-gene basis will enable the Bio-Med researcher to:

1. Perform basic analyses without the need for external resources;
2. Have immediate access to the visualization and statistics of the data;
3. Compare the expression pattern validity of different gene probes for better wet experiments decision making;
4. Improve their ability to produce significant results out of their various experiments.

COMPETING INTERESTS

The authors declare that they have no competing interests.

FUNDING

This work was funded by a NARSAD Young Investigator Grant from the Brain & Behavior Foundation, the National Institute of Psychobiology in Israel (Grant number 132-14-15) and the Leir Charitable Foundation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

We deeply thank Professor Eytan Domany for his wise advice and for providing the conditions to develop GEView. We thank Dr. David Gurwitz for providing the first dataset for testing and his wise advices.

REFERENCES

Aspland, S. E., Bendall, H. H., & Murre, C. (2001). The role of E2A-PBX1 in leukemogenesis. *Oncogene*, *20*(40), 5708–5717. doi:10.1038j.onc.1204592 PMID:11607820

Bengtsson, H. (n.d.). R.matlab. Retrieved from https://github.com/HenrikBengtsson/R.matlab

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300.

Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 9: One-way analysis of variance. *Critical Care (London, England)*, *8*(2), 130–136. doi:10.1186/cc2836 PMID:15025774

Bornhauser, B. C., Bonapace, L., Lindholm, D., Martinez, R., Cario, G., Schrappe, M., ... Bourquin, J.-P. (2007). Low-dose arsenic trioxide sensitizes glucocorticoid-resistant acute lymphoblastic leukemia cells to dexamethasone via an Akt-dependent pathway. *Blood*, *110*(6), 2084–2091. doi:10.1182/blood-2006-12-060970 PMID:17537996

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. doi:10.1038/nbt.1511 PMID:19029910

Glish, G. L., & Vachet, R. W. (2003). The basics of mass spectrometry in the twenty-first century. *Nature Reviews. Drug Discovery*, *2*(2), 140–150. doi:10.1038/nrd1011 PMID:12563305

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews. Genetics*, *17*(6), 333–351. doi:10.1038/nrg.2016.49 PMID:27184599

Howe, E. A., Sinha, R., Schlauch, D., & Quackenbush, J. (2011). RNA-Seq analysis in MeV. *Bioinformatics (Oxford, England)*, *27*(22), 3209–3210. doi:10.1093/bioinformatics/btr490 PMID:21976420

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics (Oxford, England)*, *8*(1), 118–127. doi:10.1093/biostatistics/kxj037 PMID:16632515

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *374*(2065), 20150202. doi:10.1098/rsta.2015.0202 PMID:26953178

Kallio, M. A., Tuimala, J. T., Hupponen, T., Klemelä, P., Gentile, M., Scheinin, I., ... Korpelainen, E. I. (2011). Chipster: User-friendly analysis software for microarray and other high-throughput data. *BMC Genomics*, *12*(1), 507. doi:10.1186/1471-2164-12-507 PMID:21999641

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Res., 12*(6), 996–1006. doi:10.1101/gr.229102

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics (Oxford, England)*, *28*(6), 882–883. doi:10.1093/bioinformatics/bts034 PMID:22257669

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. doi:10.118613059-014-0550-8 PMID:25516281

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47–e47. doi:10.1093/nar/gkv007 PMID:25605792

Ross, M. E., Zhou, X., Song, G., Shurtleff, S. A., Girtman, K., Williams, W. K., ... Downing, J. R. (2003). Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood*, *102*(8), 2951–2959. doi:10.1182/blood-2003-01-0338 PMID:12730115

Safran, M., Dalah, I., Alexander, J., Rosen, N., Iny Stein, T., Shmoish, M., ... Lancet, D. (2010). GeneCards Version 3: The human gene integrator. *Database : The Journal of Biological Databases and Curation*, *2010*(0), baq020. doi:10.1093/database/baq020 PMID:20689021

Schena, M., Shalon, D., Davis, R. W., & Brown, P. O. (1995). Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. *Science*, *270*(5235), 467–470. doi:10.1126cience.270.5235.467 PMID:7569999

Teuffel, O., Dettling, M., Cario, G., Stanulla, M., Schrappe, M., Bühlmann, P., ... Schäfer, B. W. (2004). Gene expression profiles and risk stratification in childhood acute lymphoblastic leukemia. *Haematologica*, *89*(7), 801–808. PMID:15257931

Tukey JW. (n.d.). Comparing individual means in the analysis of variance.

Tusher, V. G., Tibshirani, R., & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(9), 5116–5121. doi:10.1073/pnas.091062498 PMID:11309499

Ulitsky, I., Maron-Katz, A., Shavit, S., Sagir, D., Linhart, C., Elkon, R., ... Shamir, R. (2010). Expander: From expression microarrays to networks and functions. *Nature Protocols*, *5*(2), 303–322. doi:10.1038/nprot.2009.230 PMID:20134430

This research was previously published in the International Journal of Knowledge Discovery in Bioinformatics (IJKDB), 8(1); edited by Mehdi Khosrow-Pour, D.B.A.; pages 94-105; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Availability and Requirements

Project name: GEView
Project site: http://www.weizmann.ac.il/complex/compphys/software/geview/_
Programming Language: MATLAB (https://www.mathworks.com). In order to facilitate the installation step, GEView contains the following R packages: sva (Leek et al., 2012), R.matlab (Bengtsson, n.d.) and their subsidiaries.
Operating Systems: The compiled version can be launched on either Windows or Linux. Using the MATLAB source code, that can be downloaded from GEView website (http://www.weizmann.ac.il/complex/compphys/software/geview/), GEView can run on any operating system with MATLAB installed, including Mac OS.
Restrictions to use by non-academics: None
How to Cite GEView: The present paper should be cited if GEView is used in the preparation of a manuscript

Declarations

Ethics approval and consent to participate

As was stated before (see (Teuffel et al., 2004), (Bornhauser et al., 2007) for the "Zurich" dataset and (Ross et al., 2003) for the "StJude" dataset), patients who were enrolled in these studies or their guardians had given informed consent in accordance with the Declaration of Helsinki. Approval was obtained from each institutions' review board.

Consent for Publication

Not applicable

Availability of Data and Materials

The two datasets that were analyzed during the current study are available through download-links that appear in the following references: St Jude Children's Research Hospital, Memphis, TN (Ross et al., 2003) for the "StJude" dataset, and Departments of Oncology University Children's Hospital Zurich, Switzerland (Teuffel et al., 2004) for the "Zurich" dataset.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 33
An Affordable Hybrid Cloud Based Cluster for Secure Health Informatics Research

Basit Qureshi
Prince Sultan University, Saudi Arabia

ABSTRACT

This article describes how a major risk factor in the deployment of patient health records systems in the cloud is the security and privacy of data. Hybrid cloud solutions have been proposed that leverage the public and private cloud deployment to manage and alleviate accessibility, access control and privacy concerns. This article presents a privacy preserving and secure architecture for data acquisition, storage, processing and sharing. The proposed architecture is composed of a public cloud-based services that interact with a low-cost cloud computing cluster (LoC4) as a backend. A lightweight data security eco-system based on attribute based encryption is developed to provide security for public cloud-based data storage. Performance of the deployment is evaluated in a real-time deployment environment. The results show that the proposed ABE-based system is 2.3 times faster than AES-based for a variety of sizes of data blocks. It is further noted that the low-cost and affordability of LoC4 platform offers excellent opportunities for academic research in cloud based health informatics.

1. INTRODUCTION

Electronic health record (EHR) and personal health record (PHR) systems are widely available and use different technologies and standards (Bahga, 2015; Sun, 2013). The variety and size of medical health records data makes it difficult for researchers to accurately and easily integrate data from various sources. Due to the high cost of building and maintaining specialized data centers many healthcare providers have been outsourcing PHR systems to third part cloud service providers such as Microsoft Health Vault, Google Health, General Electric's Centricity Patient Online, eClinicalWorks and many more. Cloud computing provides storage for large scale data on external servers allowing researchers and develop-

DOI: 10.4018/978-1-7998-1204-3.ch033

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ers to easily access the data (Fang, 2016). Although the technology is being used in recent years with many benefits such as reducing the cost for maintaining servers whereas improving the availability of the systems, there have been wide spread privacy concerns since personal healthcare information can be exposed to unauthorized parties.

Due to the sensitivity of personal information, the cloud based PHRs could potentially become targets of various malicious behaviors. Furthermore, Guo et. al. in (Guo, 2016) and (Abbas, 2017) describe various security issues and threats in cloud computing environment. In (Asija R., 2016), researchers develop a security aware public cloud based SaaS model for healthcare applications embedding XML based meta data for PHRs. Works presented in (ZhouJ, 2015), (ZhouCao, 2015) and (Li, 2013) have proposed several cloud-based secure systems, especially cloud-based PHR systems using encryption where PHR data is encrypted and stored in cloud based storage. In (Man, 2017), (Wang, 2014) and (Zhou, 2015) authors deduce that any public cloud based security model may have inherent threats of trust, security and privacy. They highlight that the complexity of key management in encrypted PHR systems increases for the owners and users resulting in Quality of Service issues. (Liu, 2017) note that hybrid cloud environment integrating the public and private cloud infrastructure is a more applicable solution for PHR sharing.

In this paper, we take inspiration from (Liu, 2017) to investigate requirements for the deployment of a secure hybrid cloud environment for a PHR system. A two-tier hybrid cloud architecture is proposed addressing security and privacy preserving mechanisms for information storage and retrieval in the cloud. A light weight data security eco system is developed to provide Attribute Based Encryption (ABE) for public cloud based data storage that enables multiple users to share data. We further introduce LoC4, an affordable and low-cost cloud computing cluster for health informatics research in universities and academic institutions. The proposed hybrid cloud based PHR system is developed and deployed using an experimental setup coupling public and private cloud based implementations. We conduct several experiments to analyze the performance criteria including computation efficiency, storage and network efficiency, task completion rate, response time etc. Furthermore, we compare the performance the proposed Encryption scheme and compare the response time for task execution for various sizes of data blocks. Results show that LoC4 provides an affordable hybrid cloud based backend solution for low cost, energy efficient and secure deployment.

The contributions of this paper are in the following:

- Requirements of a secure hybrid cloud environment for a PHR system are presented. Based on an extensive investigation, a three-layered architecture is proposed including: i) secure data acquisition, ii) storage of PHRs in public cloud and iii) encryption and privacy layers;
- A two-tier, light weight encryption eco-system is developed encompassing public cloud service and private cloud based cluster deployment;
- Design details and implementation of a LoC4 cluster using SBCs is provided. The Cluster deployment parameters are optimized for performance and energy efficiency;
- A light weight middleware for communication and interaction between public cloud service and the LoC4 Cluster is developed. This middleware enables real time performance measurement of the proposed system.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The remainder of this paper is organized as follows: Section 2 provides description of requirement specifications for hybrid cloud architecture for PHR systems. Section 3 details the two-tier light weight eco-system. Section 4 provides design details for the LoC4 cluster as well as performance specific variables for deployment of Apache Hadoop on the cluster. Performance evaluation of the proposed system is presented in section 5. Section 6 presents related works with conclusions and future works presented in section 7.

2. REQUIREMENTS SPECIFICATIONS FOR A CLOUD BASED PHR SYSTEM

A hybrid cloud based PHR system is composed of various subsystems interacting in a public and private cloud environment. Figure 1 shows various parts of a hybrid cloud deployment of a PHR systems for a Healthcare provider (HCP). Data deemed secure and private is kept in a data center composed or various clusters whereas unsecure data or data with no privacy concerns is kept for public access in a public cloud. In this section, we provide requirements specifications for a secure hybrid cloud based PHR system.

2.1. Storage in Cloud Based PHR System

A cloud based PHR system needs to address secure data acquisition and storage requirements before further processing of data. Figure 2 shows an architectural map of the three layers of data flow in a PHR system. In what follows, we describe requirements for these data flows.

At the acquisition stage, data is collected using various mechanisms for data entry. It could be as complex as a Wireless Body Area Network (WBAN) (Page, 2014) (Page, 2015) consisting of various wireless wearable sensors for specific medical applications such as blood pressure, thermometer etc. These sensors and devices can connect to a data collection device using Bluetooth, Zigbee protocols or similar wireless communication medium, and transmit the information to an intermediate cloudlet for synchronizing and processing information (Soyata, 2016). This data could be acquired, stored on the cloudlet or transmitted on demand for access by the HCPs.

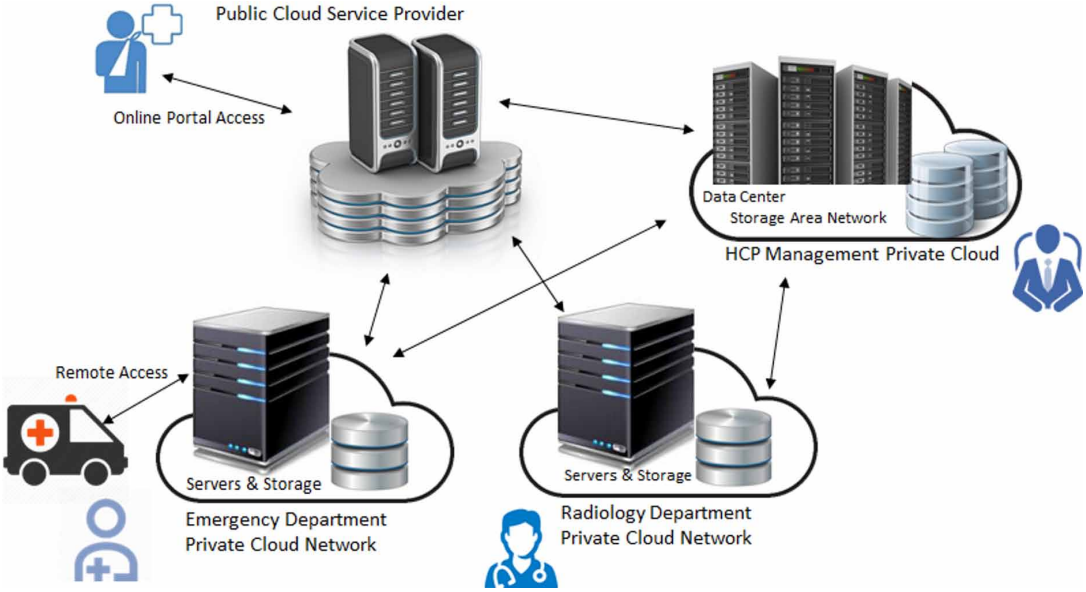
Reliable, scalable and secure storage in the cloud is perhaps the most important functionality of Cloud based PHR systems. Storage services based on public clouds such as Microsoft's Azure storage service and Amazon's S3 provide customers with scalable and dynamic storage. These services allow users to move their data into the cloud and avoid costs of building and maintaining a private storage infrastructure using pay as you go cost model. This offers various benefits such as availability (i.e., being able to access data from anywhere) and reliability (i.e., not having to worry about backups) at a relatively low cost. Consumers can trade privacy for convenience of software services, however enterprises and government agencies are more concerned with the preservation of confidentiality and integrity of data.

2.2. Data Privacy Requirements

The Health Insurance Portability and Accountability Act (HIPAA) (HIPAA US, 2006) states that data privacy must be protected within every layer of a PHR system. Here we describe the requirements for each layer of the system:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A hybrid cloud environment for a PHR system at HCP



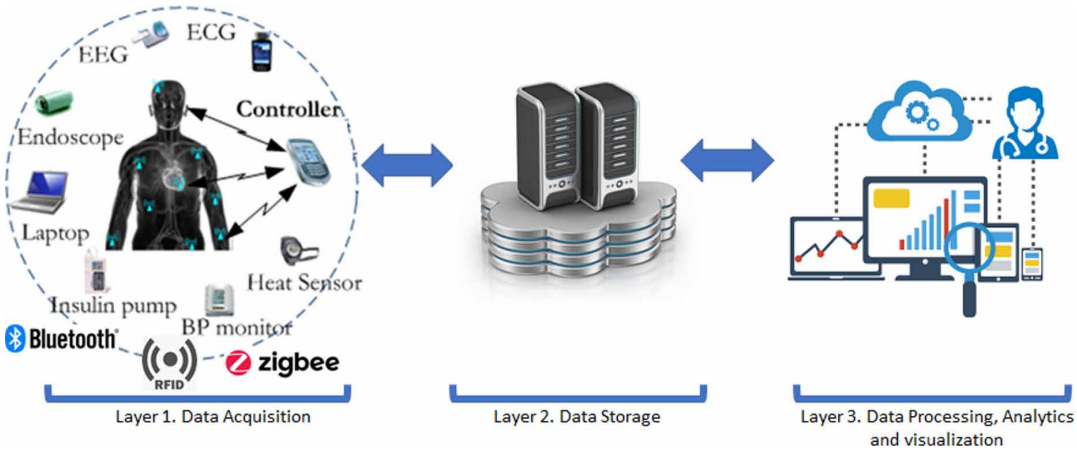
1. **Data acquisition Layer:** The acquisition layer in Figure 2 is composed of Wireless Body Area Network sensor devices. Typically, these devices have limited computational capacities and operate on finite battery power. Encryption schemes used to protect the communication within BAN sensors and BAN-to cloudlet communications should not be computationally intensive. (Page A., 2014) present a privacy preserving mechanism for secure information exchange using Advanced Encryption Standard (AES);
2. **Data Storage Layer:** There can be various stakeholders involved in storage, retrieval and sharing of data from the system including, patient, doctor/nurse, emergency unit staff etc. Allowing Cloud service providers to manage sensitive information such as EHR could potentially raise security issues. An untrustworthy Cloud service provider may leak PHR to medical companies for profit. It is suggested to encrypt all the data placed in the Cloud, however, this could have a severely negative affect on the Quality of Service as the volume of data grows. Furthermore, the cost of encryption also increases linearly in time and space.

Another issue is the QoS in searching and indexing of data. To enable searching over the data, a user must either store an index locally, or download all the (encrypted) data, decrypt it and search locally. Both approaches negate the benefit of cloud computing and are therefore poor choice for processing of large scale data thus increasing the overhead of security and volume of communications.

Conventional encryption techniques such as symmetric key encryption does not work as it requires both parties to use a symmetric key. Asymmetric encryption techniques such as Identity based encryption (IBE) may work, but it requires one party to decrypt, other than this party, no one can decrypt the data. It can be suitable in a scenario when a data record is shared only between two parties, however for multiple actors such as doctors, nurses, insurance companies, patients, etc., management becomes an overhead:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Architectural layout of cloud based PHR system requirements



3. **Data Processing, Analytics and Visualization Layer:** To process data, it must first be decrypted. Decryption of this data in the intermediate stage requires trusted storage infrastructure possibly at the site of the HCP necessitating a private cloud infrastructure. After the data has been processed, it needs to be removed from the intermediate storage. Furthermore, this encryption and decryption of data adds to overall overhead of computation, processing and visualization of data. Efficiency, reliability and integrity must be preserved for data processed in the system.

It is imperative that in an enterprise environment, many stakeholders may have access to various kinds of data. Doctors, patients, nurses, emergency unit staff, medical insurance personnel etc. to name a few, require strict access control mechanisms in place. Furthermore, in WBAN environment, various devices could be owned by a consumer or even shared by others. For any efficient solution, it is crucial to address the above-mentioned issues thoroughly and effectively.

3. SECURE HYBRID CLOUD ARCHITECTURE FOR PHR SYSTEMS

In this section, we present the design and architecture for a secure hybrid cloud based PHR system. The proposed architecture considers the requirements of data acquisition, storage and analysis while preserving the data privacy requirements. We present a high-level architecture of the proposed system in Figure 3 which incorporates various components into public and private cloud tiers. All services interact together to provide secure, reliable and scalable solution.

3.1. Public Cloud Service

The public cloud service consists of various components including encrypted data storage, open access data storage, Health analytics engine, Audit Logs, Enterprise Access control list and Key verification module for data integrity. We adopt attribute-based encryption (ABE) as the cryptographic primitive to serve this purpose. By using ABE, access control definitions express the user attributes (meta-data) or data (PHR) which enable a user to selectively share their data without knowledge of the other party's identity.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

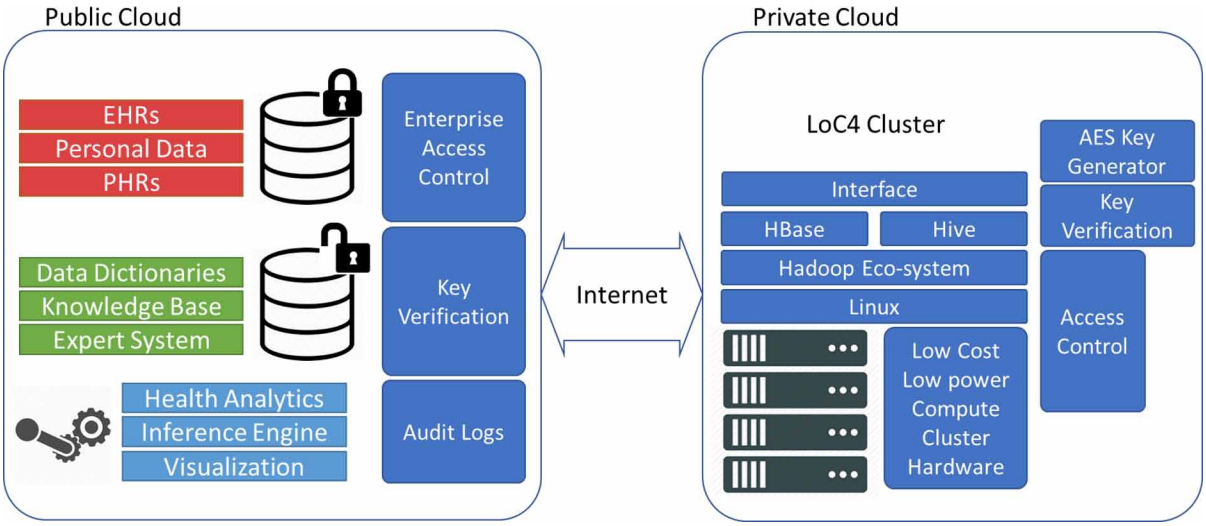
Privacy critical data such as personal health records need to be secured on the public cloud, whereas data with low privacy requirements such as data dictionaries etc. or intermediary data such as data generated from health analytics or visualization, many not require stringent security measures. It is important to classify these kinds of data to improve the efficiency of the system by leveraging the amount of computations and time consumed in communication. Large volumes of data generated at the Data acquisition layer needs to be stored and shared with appropriate consumers with valid credentials. Access Control parameters are defined to identify appropriate levels of users mapped with data in storage. This mechanism allows stakeholders with appropriate credentials to access and share data. Access Control parameters are defined in the private cloud service and are periodically updated in the public service.

We develop a custom encryption approach that is efficient and reliable and is described further in this section. Audit logs are stored for all transactions carried out by various consumers. These logs are beneficial for forensics in determining illegal behavior. The public cloud service is implemented in an open access system and deployed on a public cloud service provider.

3.2. Private Cloud Service

The private cloud service essentially hosts a cluster of machines implementing a small data center at the premises of the enterprise. The infrastructure is secured in the enterprise locale with physical, data and network security measures in place. The data center hosts low cost, efficient and eco-friendly LoC4 cluster. The machines in the LoC4 cluster run open source operating system such as Linux with Hadoop eco-system deployment for distributed task processing and management. Access control information for the enterprise is stored in HBase and Hive data warehouse which are encrypted using commonly available encryption techniques. Additional modules for encryption key generation and verification are implemented in the private cloud. These are necessary for efficient and reliable storage and sharing of data. Communication takes place through the interface which implements a light weight communication protocol for data transmission between the public and private cloud services. Section 4 details the design and implementation of the LoC4 cluster which is deployed using low cost low power Single Board Computers (SBCs).

Figure 3. High-level architecture



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.3. Security and Privacy

Here we address the security implications and proposed solutions at the various layers of data in a PHR system.

3.3.1. Data Acquisition Layer

Data acquisition devices usually have limited capabilities for storage or processing large scales of data. This implies that any encryption schemes used for security in data acquisition must not be computationally intensive. An obvious solution is to use the Zigbee protocol which is based on the AES encryption scheme and can easily be implemented. Communicating devices must agree on a secret-key before using AES encryption by using generic key exchange algorithm such as Diffie-Hellman (DH) (Diffie, 2006). Irrespective of the kind of encryption scheme, a consumer must agree on key(s) to encrypt/decrypt messages with a service provider. Advanced Encryption Standard (AES) is one of the most widely used symmetric key encryption algorithms and is accepted as an industry and a government applications standard. AES is optimized for speed, low memory footprint and energy efficiency. Its low resource intensity allows AES to run on a wide range of hardware platforms. Communication of devices can be also secured by using biomedical signals as noticed in (Venkatasubramanian, 2010).

3.3.2. Secure Data Storage, Sharing and Processing

We use a variant of Attribute based encryption (ABE) scheme. Attributes set is defined where each member of this set is given associated private keys which are generated by third parties. The HCP provides encryption with a token, any user with appropriate attribute set can decrypt the data as long as it fulfils the token. The token can be generated such that the retrieval procedure reveals nothing about the files or the keywords (meta-data) except that the files contain a keyword in common. This light weight technique addresses the inefficiency of downloading an entire encrypted block and processing only a limited part of it. Furthermore, the coupling of access control mechanisms (attributes) allows reliable sharing of appropriate data with proper stakeholders.

The data is encrypted using an access control policy based on credentials. Only the user whose credentials satisfy the access policy can have access to the data. The attributes can be the profession (e.g., Doctor, Nurse) or the department (e.g., Cardiology, Emergency) of the user. An access policy can be defined as joints and disjoints with various attributes such as:

(Doctor AND Pediatrics) OR (Nurse OR Intensive-Care-Unit)

Which gives access to a Doctor in Pediatrics Department or a Nurse or an Intensive-Care-Unit staff member.

We present a publisher/subscriber model for the encryption and decryption of Data Blocks. Two algorithms are defined PushDataBlock (Algorithm 1) and PullDataBlock (Algorithm 2), that control read / write of data blocks using a combination of key and access control meta data, in the public cloud service. In the following text, we describe the working of these two algorithms to enable secure data communication and storage of a data block in the public cloud service.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 1. PushDataBlock

```
input: DataBlock B_i, SecretKey S_k, AccessControlMetaData AC_MDi
output: Encrypted DataBlock EB_i
  for i = 1 to n - 1 do
     KeyGenerator(B_i, S_k, AC_MDi)
     PushCloud(EB_i)
  end
```

Algorithm 2. PullDataBlock

```
input: SecretKey S_k, Keyword
output: Decrypted DataBlock DB_i
  for i = 1 to n - 1 do
     RequestToken(AC_MD, S_k, Keyword)
     AcquireToken(S_k)
     PushToken(token, S_k)
     PullDataBock(DB_i)
     VerifyDataBlock(DB_i)
 end
```

Each subscriber is allocated a Key S_k to facilitate access to the private cloud service, these keys are managed by the administrators at the enterprise and assigned to stakeholders with appropriate access control credentials.

Assuming a data block B_i was generated (data acquired from WBAN cloudlet, etc.) consisting of one or many records / files and is ready to be uploaded to the public cloud storage, the publisher would call the *PushDataBlock* with an access control meta data AC_{MDi} for this data block B_i.

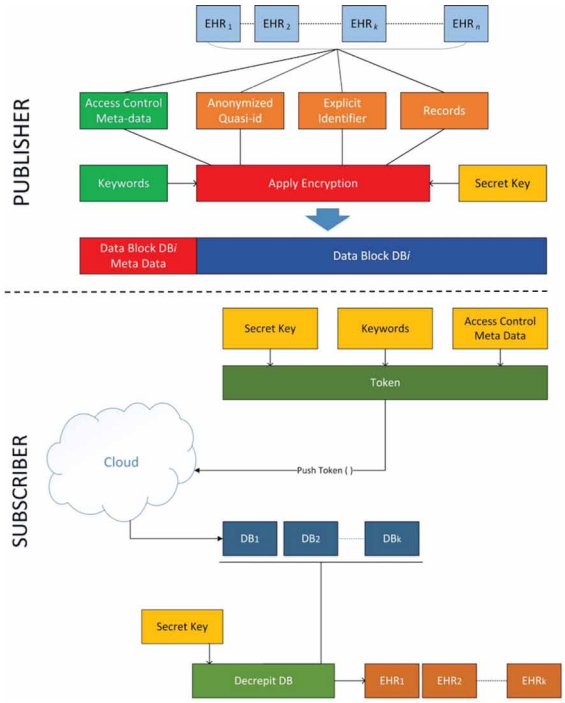
The *KeyGenerator* is responsible for properly indexing and encrypting the data block. KG is responsible for encrypting various data blocks using access control information provided and appends metadata for identification of datablocks such as timestamp, size, keywords, ownership etc. It encrypts the data block in such a way that given a *token*, a subscriber can retrieve pointers to the encrypted records/files that contain a *keyword*. Without appropriate *token*, the data block would not be decrypted. The tokens cannot be generated without a secret key S_k. It must be noted that the retrieval procedure does not reveal any details within the data block except only the searchable keywords. Once the *KeyGenerator* has encrypted the DataBlock B_i, *PushCloud*(B_i) is called to push the data block to the public cloud storage.

When a subscriber intends to access a data block for processing or sharing, it needs to call the *PullDataBlock* with the secret key S_k, and the access control meta data AC_{MD}. *RequestToken* call is made for a search keyword term to the private cloud service. Based on the access control policy, the *AcquireToken*(S_k, AC_{MD}) generates a token which is then sent to the public cloud service using the *PushToken*(*token*). The public cloud service uses the token to find the appropriate encrypted documents. The *PullDataBlock* call returns the selected documents as data blocks to the subscriber. At any point and time, the subscriber can verify the integrity of the data blocks by calling the *VerifyDataBlock* call.

Figure 4, shows the publisher subscriber model used for encryption of Data Blocks. In the next sections, we present details for deployment of the proposed model on a LoC4 cluster.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Publisher subscriber model for encryption of data blocks in the framework



4. DESIGN OF LOC4 CLUSTER

This section presents the architecture and configuration of the LoC4 cluster deployed in this experimental study.

4.1. Single Board Computers

A single-board computer (SBC) is a complete computer built on a single circuit board (Baun, 2016). A SBC incorporates microprocessor(s), memory, I/O as well as host of other features required by a functional computer. A wide range of programming languages are supported by these platforms. In this work, we investigate use of Raspberry Pi Model 3B (RPi, 2017) as well as Hard Kernel Odroid Xu-4 (Odroid, 2017). A summary of various features of these computers can be seen in Table 1.

4.2. LoC4 Cluster Deployment

This section presents the architecture and configuration of the LoC4 cluster deployed in this experimental study.

4.2.1. Design of the Cluster

The LoC4 cluster is composed of 11 SBCs of which five are RPi Model 3B computers in addition to six Odroid Xu-4 computers interconnected with power supplies, network cables, storage modules, connec-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Comparison of Raspberry Pi 3B and Hard Kernel Odroid computers

	Raspberry Pi Model 3B	Odroid XU-4
Processor (CPU)	1.2 GHz Quad core 64-bit ARM Cortex-A53	Samsung Exynos5 Octa ARM Cortex-A15 (@ 2.0 GHz) and Cortex-A7 (@1.3GHz) CPUs
GPU	400MHz Video Core IV multimedia	Mali T628 Open GL 3.0
Onboard RAM	1 GB LPDDR2-900 SDRAM (at 900MHz)	2GB LPDDR3 at 933MHz
Ethernet / Network	10/100 MB Ethernet RJ45 Jack. Onboard 2.4GHz Wi-Fi 802.11n and Bluetooth 4.1	10/100/1000 MB Ethernet RJ45 Jack
Storage	Micro SD Card	Micro SD Card and eMMC 5.0 flash storage
Audio/Video	Micro HDMI	HDMI (standard) supports 1080p video
Power Consumption	2.0 W (idle) 3.8 W (under load)	2.5 W (idle) 4.5 W (under load)
USB Ports	4x USB 2.0 ports	1x USB 2.0, 2x USB 3.0
Released	2016	2015
Price (US$)	35 $	79 $

tors and cases. All the Raspberry Pi computers are equipped with 16 GB Class-10 SD cards for primary bootable storage. The Odroid Xu-4 devices are equipped with 32GB eMMCv5.0 modules. The Odroid Xu-4s are housed in a compact layout racks using M2/M3 spacers, nuts and screws. The Raspberry Pi 3B's are housed in cases and connected to the rest of the cluster. Currently each Raspberry Pi computer is individually supplied by 2.5Amp power supply; each Odroid Xu-4 computer is supplied by a 4.0Amp power supply that provides ample power for running each node. Each SBC's network interface relates to a Cat6e Ethernet cable through the RJ-45 Ethernet connector. All Ethernet cables connect to a 16-port Cisco switch which in turn is connected to the university network equipment. The size of the cluster can be easily scaled by introducing a core switch that connects to the 16-port switch. Additional nodes can be included in the network configuration scaling the size of the cluster.

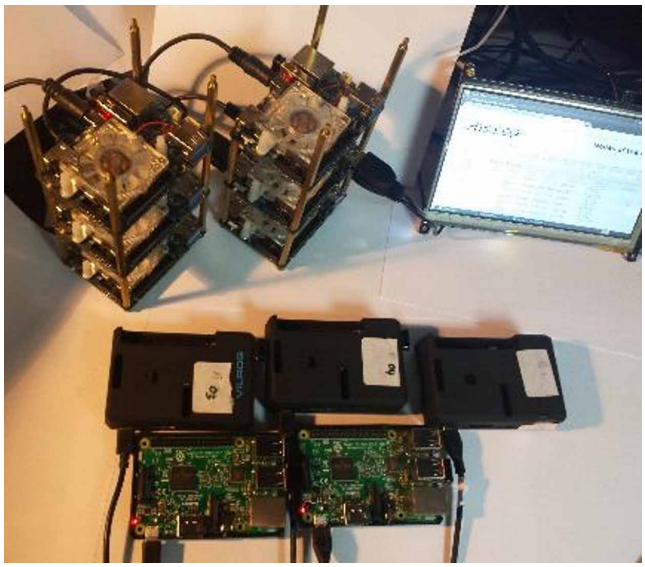
For Hadoop deployment, one of the Xu-4 SBC is used as the master node whereas the rest of the nodes server as slaves' nodes. We compare the performance of Xu-4 sub cluster (composed of five Xu-4 nodes) with RPi sub cluster also composed of five RPi 3B nodes. Figure 5 shows the layout of the LoC4 cluster. The purchase cost for all equipment for the LoC4 Cluster was $1179.

4.2.2. Hadoop Deployment

Apache Hadoop (Hadoop, 2014) is an open source framework that provides distributed processing of large amounts of data in a datacenter. Hadoop version 2.6.2 was installed due to availability of YARN daemon which improves the performance of the map-reduce jobs in the cluster. To optimize the performance of these Clusters, yarn-site.xml and Mapred-site.xml were configured with 512 MB of resource size allocation. The primary reason for this setting is the limited amount of RAM available in the RPi Model 3B. The Raspbian operating system as well as the shared GPU memory bus consume over 200MB or RAM out of a total of 1GB. The default container size on the Hadoop Distributed File System (HDFS) is 128 MB. Each SBC node was assigned a static IPv4 address based on the configuration and all slave nodes were registered in the Master node. Table 2 provides detail of important configuration properties for the Hadoop environment.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The LoC4 cluster with 6 Xu4 (rack mounted) and 5 RPi 3B in cases



For YARN Resource manager, we allocated up to 4 cores which means up to 4 containers can execute per node (one container per core). The replication factor for HDFS is 2 which means only two copies of each block would be kept on the file system. The LoC4 Cluster was tested extensively for performance using TeraSort benchmark.

The LoC4 cluster was connected to the Internet using a static IP address for experimental evaluation presented in the next sections.

5. PERFORMANCE EVALUATION

In this section, we present a performance evaluation study of the proposed Hybrid Cloud based PHR system. We first look at the performance evaluation of the LoC4 cluster in terms of processing speed, storage read/write and networking. The objective of this study is to highlight the intrinsic capabilities of SBCs as well as to showcase the performance of these SBC when deployed as a cluster. We use popular benchmarking techniques widely used in the literature to highlight the capabilities of LoC4. Second, we present performance in terms of Response time for the *PushDataBlock* and *PullDataBlock* algorithms presented in section 3.

5.1. Evaluation of LoC4

5.1.1. CPU Execution Times

The benchmark suite Sysbench[1] was used to measure the CPU performance. We execute the Sysbench benchmark[2], testing each number up to value 10,000, if it is a prime number, for *n* number of threads. Since each computer has a quad core processor we run the sysbench CPU test for 1, 2, 4, 8 and 16

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Hadoop configuration properties for LoC4 cluster

mapred-site.xml Properties	Value
yarn.app.mapreduce.am.resource.mb	512
mapreduce.map.cpu.vcores	1
mapreduce.reduce.cpu.vcores	1
mapreduce.map.memory.mb	512
mapreduce.reduce.memory.mb	512
mapreduce.input.fileinputformat.split.minsize	8 MB
YARN-site.xml Properties	
yarn.nodemanager.resource.memory-mb	1024
yarn.nodemanager.resource.cpu-vcores	1
yarn.scheduler.minimum-allocation-mb	256
yarn.scheduler.maximum-allocation-mb	1024
yarn.scheduler.minimum-allocation-vcores	1
yarn.scheduler.maximum-allocation-vcores	4
yarn.nodemanager.vmem-pmem-ratio	2
hdfs-site.xml Properties	
dfs.replication	2

threads. We measure the performance for this benchmark test for Raspberry Pi Model 2B, Odroid Xu-4 as well as an Intel i7 3.0 GHz 4th Generation personal computer (for comparison purposes). The CPU execution times scale well with increased number of threads. Sysbench test runs with n = 2 and n = 4 threads significantly improves the execution times performance for all processors by 50%. With n = 8 and n = 16 threads the test results yields almost similar execution times with little improvement in performance. Table 3 shows results for the Sysbench. The execution times for Odroid Xu-4 are 6 times better as compared to Raspberry Pi Model 3B. The increased number of threads (larger than 8) does not provide gain in performance of Odroid Xu-4 over Raspberry Pi. It is also noted that the execution time for Raspberry Pi is further extended with larger *n*.

5.1.2. Storage Performance

Poor storage read/write performance can be a bottleneck in clusters. Compared to a server machines, an SBC is handicapped in terms of availability of limited storage options. In our experimentation, the Raspberry Pi's were equipped with 16GB SanDisk Class 10 SD Cards, whereas the XU-4 devices were equipped with 32GB eMMC memory cards. Both memory cards were loaded with bootable Linux distributions.

FIO[3] was used for benchmarking of random read and write with various block sizes. We consider the throughput with 8 threads each working with a file of size 512MB with a total 4GB of data. These parameters were set specifically to avoid buffering and caching in RAM issues which is managed by the underlying operating systems that can distort the results; i.e. the data size (4GB) selected is larger than the onboard RAM available on these devices. Table 4 shows the comparison of buffered and non-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

buffered random read and write from both devices with block size 8KB. As can be seen from Table 4, the read throughput (buffered) of Odroid with eMMC memory is at least twice as fast as the Class 10 SD Card on the Raspberry Pi whereas the non-buffered read is more than three times better. Similarly, for buffered write operations, Odroid Xu-4 with eMMC module throughput is more than twice better when compared to the Class 10 SD Card in Raspberry Pi.

5.1.3. Network Performance

was measured using the iperf 4v3.13 with the NetPIPE5 benchmark version 3.7.2. Through various sets of runs, iperf states the network throughput to be 82-88Mbits per second for the RPi as well as the Xu-4 SBCs. The NPtcp, NetPIPE benchmark using TCP protocol, involves running transmitter and receiver on two nodes in the cluster. In our experimentation, we executed the receiver on the LoC4 master node (Xu-4 device) with 1000 KB as maximum transmission buffer size for a period of 240 milliseconds. The transmitter was executed on the individual SBCs one by one. As can be seen from Figure 6, the network latency for both devices with small payload is almost similar. As the payload increases we observe slight increase in network latency between both sets of devices. On the other hand, we observe a spike in throughput with message size 1000 bytes. This proves that network layer communication overhead dominates the overall transfer time for small payloads. For larger payloads, the communication rate correlates with data rate at the network link, utilization of the communication medium at the time or the traffic at the network switch. Contrasting the performance of Xu-4 and RPi SBCs we note the visible difference in throughput between the two, which is due to the poor overall Ethernet performance of the Raspberry Pi.

Table 3. The average CPU execution time for nodes with n threads

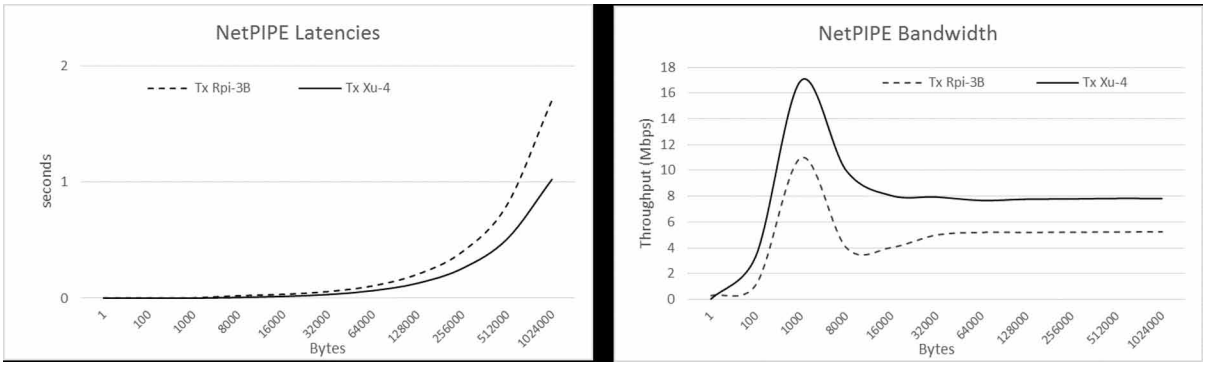
	CPU Cores	Clock Rate GHz	CPU Execution Time with *n* Threads				
			1	2	4	8	16
Raspberry Pi 3B	4	1.2	382.2	192.07	96.31	98.69	99.73
Odroid Xu-4	8	2.0 and 1.4	83.38	41.68	25.336	17.66	18.02
Intel i7 4th Gen	4	3.0	8.508	4.272	2.228	2.27	2.23

Table 4. Read and write throughput (KB/sec) for individual devices in the clusters using FIO

	Read Throughput (KB/second)		Write Throughput (KB/second)	
	Buffered	Non-Buffered	Buffered	Non-Buffered
Raspberry Pi 3B with 16 GB Class 10 SanDisk SD Card	3,688	2,341	1,318	1,267
Odroid Xu-4 with 32GB eMMCv5.0 Module	9,197	6,883	3,220	2,519

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. NetPIPE benchmark results for Xu4 and RPi devices considering latencies and bandwidth with data size in terms of bytes on the x-axis



5.1.4. Hadoop TeraSort Benchmark

Hadoop provides various benchmarks for performance of the Hadoop cluster. TeraSort is used in this study to observe the performance of LoC4 cluster in terms of task execution time. TeraSort combines testing the HDFS and MapReduce layers of a Hadoop cluster and consists of three MapReduce programs, TeraGen, TeraSort and TeraValidate. TeraGen is used to generate large amounts of data blocks, consequently, TeraGen is a write intensive I/O benchmark. TeraSort generates set of sample keys by sampling the input data generated by TeraGen before the job is submitted and writes the list of keys into HDFS. The input and output format, which are used by all three MapReduce programs, reads and writes the text files in the correct format. The TeraSort benchmark is CPU bound during the map phase and I/O bound during the reduce phase. TeraValidate verifies that the output of the TeraSort is globally sorted. We run TeraGen, and TeraSort on LoC4 cluster with 400MB, 800MB, 1,600MB and 3,200MB data sizes.

We observe the job execution time for each run for comparison and analyze the performance on the cluster. Since TeraGen is I/O intensive, the write speeds of the memories/storage in corresponding nodes in the clusters plays a major role in degrading the overall job completion time. The average job completion times for RPi and Xu-4 sub-clusters can be seen in Figure 7. The completion time for TeraGen on Xu-4 and RPi is correlating for various data sizes with Xu-4 being 3 times faster compared to RPi. We compute the execution time for TeraSort. For all experiments, we use the same number of map and reduce tasks on each cluster. The TeraSort benchmark is CPU bound during the map phase, i.e. reading input files and sorting tasks are carried out whereas it is I/O bound during the reduce phase, i.e. writing output files in the HDFS. We observed that 33-40% of job completion time occurred in map phase while 49% or more time was spending in reduce tasks overall for the majority of TeraSort jobs run on the cluster. We observe that as the data size increases the execution time also increases for RPi. Figure 7 also shows the comparison of the ratio of execution time degradations for RPi against Xu-4 sub-cluster. For smaller data size of 400MB the Xu-4 is 3 times faster whereas for larger data size of 3.2 GB, the Xu-4 is 3.4 times faster.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5.2. Evaluation of Hybrid Cloud Based PHR

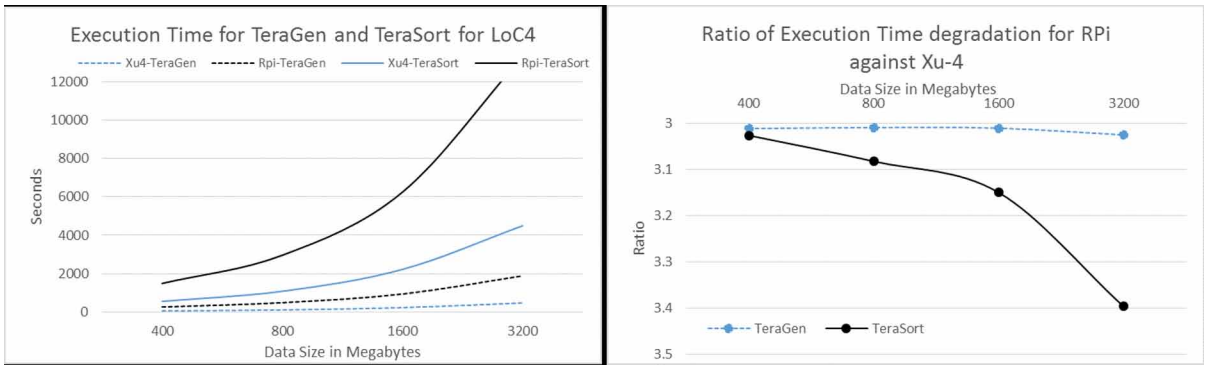
In this section, we evaluate the performance of the Hybrid Cloud based PHR system presented in section 3, in two phases. First, we write a Linux based script to encrypt data blocks of various sizes using AES and push these to a public cloud service and observe response time. Next, we compare the response time with the ABE based *PushDataBlock* encryption algorithm to determine the cost of encryption in terms of response time. Second, we measure response time at the public cloud service by executing the subscribe modules. The data blocks are pulled from the cloud based on user provided attributes. For this experimental study, xml files of various sizes are used to mimic datablocks. Each file is assigned with access control information and useful meta-data such as filename, keywords, etc.

5.2.1. Response Time for Publisher

The purpose of this set of experiments is to validate the functioning of all modules for publisher access in the Hybrid Cloud based PHR System. A Ubuntu based virtual machine instance with 4 CPU cores, 2GB RAM and 50GB storage is launched in DreamHost. A standalone Hadoop 2.6.2 installation is initiated on this virtual machine. Various sets of xml files are generated with different sizes and are pushed into a HDFS directory from the master node in the LoC4 cluster. These files are encrypted using Advanced Encryption Standard (AES) with 128-bit length keys. We measure the response time in terms of task completion time, at the master node in the LoC4 cluster.

Next, we write a light weight middleware using bash script which is responsible for executing the publisher components of the system. The *KeyGenerator* Module is called to encrypt xml files of various sizes (1.8 MB to 394MB) using the access control parameters including up to 40 attributes and meta data for files. Access control parameters were generated for four users (Doctor, Nurse, Manager and Patient). These files were then pushed to the public cloud using the *PushDataBlock*. Again, we measure the response time in terms of task completion time at the LoC4 cluster master node. Table 5 shows the comparison of response time (time to completion) of both sets of experiments. Results show that the proposed scheme using ABE is on average 2.3 times faster than the AES based scheme in terms of response time for task completion. Figure 8 shows the response time for execution of the AES based technique versus the proposed mechanism. It must be noted however, as the Data block size increases the task completion time for both also increase linearly.

Figure 7. TeraGen and TeraSort execution time for LoC4 cluster



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5.2.2. Response Time for Subscriber

These experiments validate the subscriber modules of the proposed system, namely, *KeyGenerator*, *PullDataBlock*, *RequestToken* and *AcquireToken*. Bash script is written in the DreamHost based virtual machine requesting data blocks based on a query. Given the access control parameters and the search keywords as part of the query, *RequestToken* sends a request to the LoC4 Cluster. The LoC4 *AcquireToken* provides a token which is sent to the public cloud service. Using this token, *PullDataBlock* decrypts the block that can be accessed by the user on the public cloud.

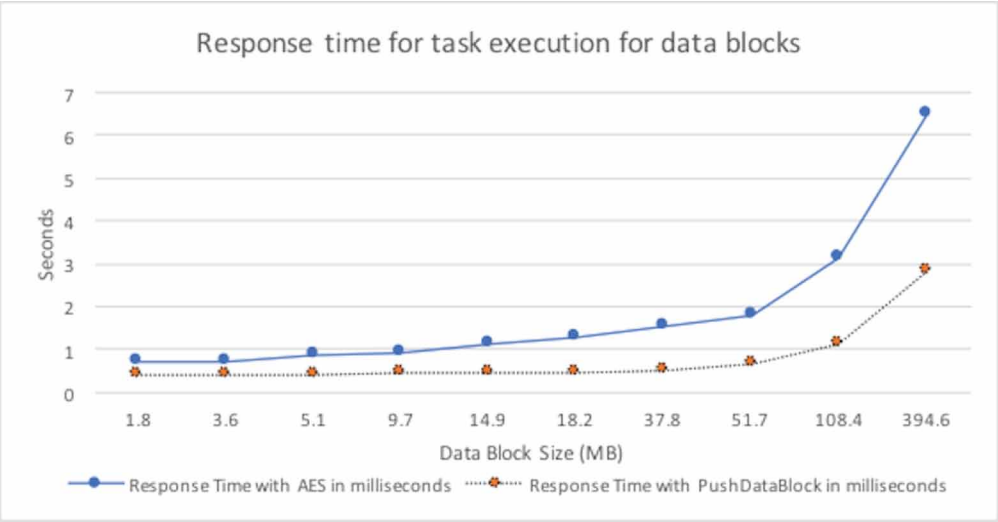
Our results show that the subscriber is successfully implemented with reasonable response times. It must be noted due to the settings in the SBC based LoC4 cluster, the minimum response time for executing Hadoop jobs is approximately 3 seconds. Given these conditions the response time for smaller workloads i.e. less than 50MB for the proposed system is within two seconds. It is observed, as the size of the *DataBlocks* increases, the time to completion also increases linearly.

6. RELATED WORKS

In recent times, many cloud service providers such as Google, Microsoft and Amazon have provided PHR services through the cloud. A PHR inherently saves privacy data of individuals which leads to the need for prevention of this data from unauthorized access. Gou et.al in (Gou, 2016) present a survey on various security issues and challenges in cloud computing environment.

In (Asija, 2016), researchers develop a security aware public cloud based SaaS model for healthcare applications embedding XML based meta data for PHRs. Work detailed in (ZhouCao, 2015) presents a privacy preserving protocol for cloud assisted e-healthcare data sharing. Authors in (Li, 2013) proposed a cloud-based PHR systems framework using encryption where PHR data is encrypted and stored in cloud based storage. (Qureshi, 2014) provide a cloud based framework for big data applications in healthcare.

Figure 8. Response time for task execution of data blocks using AES and proposed schemes



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Response time for various publisher access tasks

DataBlock(s) Size in MB	Response Time With AES in Milliseconds	Number of Attributes	Response Time With *PushDataBlock* in Milliseconds
1.8	7098.2	20	4009.9
3.6	7211.9	20	4101.2
5.1	8633.6	20	4208.1
9.7	9221.1	20	4304.8
14.9	11013.9	20	4408.0
18.2	13009.8	40	4610.0
37.8	15291.3	40	5022.6
51.7	18050.2	40	6814.0
108.4	31167.0	40	11089.1
394.6	64985.2	40	28223.1

The proposed framework leverages the benefits of cloud computing for addressing many issues in traditional PHR systems. A proof of concept is provided although no details of implementation or deployment of the framework are available. Basu et. al in (Basu, 2012) present Fusion, an experimental open, cloud-based platform for large-scale, low-cost delivery of healthcare applications. Fusion addressed use of patient-centric management of electronic health records and facilitates the secure and seamless sharing of EHRs among stakeholders within a healthcare system. (Peyton, 2017) provide a Health informatics system design approach to support community based care delivery.

Authors in (Wan, 2013) present a detailed study on cloud-enabled architecture and its applications in pervasive healthcare systems. They provide a prototype implementation of the proposed mobile healthcare system. Wang et.al. in (Wang, 2014) use a case study of mobile-cloud based electrocardiograph monitoring and analysis and develop a hybrid mobile-cloud prototype. The experimental results show that the proposed approach can enhance the conventional mobile-based medical monitoring. (Alghabban, 2017) develop a cloud based mobile healthcare framework for dyslexic students. The proposed tool was used as a e-learning tool to enhance students learning capabilities. Nepal et.al in (Nepal, 2015) address the requirements for deploying secure hybrid clouds for Healthcare data. They stress on the need for encryption of PHR in public cloud environments and availability of secure development APIs.

In (Man HoAu, 2017) authors deduce that any public cloud based security model may have inherent threats of trust, security and privacy. They highlight that the complexity of key management in encrypted PHR systems increases for the owners and users resulting in Quality of Service issues. (Zhou, 2015) note that hybrid cloud environment integrating the public and private cloud infrastructure is a more applicable solution for PHR sharing. (Liu, 2017) present a PHR access control scheme allowing access policies to be embedded into PHRs. They conduct simulations studies to verify that the proposed scheme is suitable for mobile-cloud health care. (Zhang, 2017) provide a Quality of Service perspective in Cloud services computing and provide a systematic and practical overview of Quality of Service prediction in cloud and service computing.

Most existing work has adopted ABE (Ostrovsky, 2007) as the cryptographic tool to achieve fine-grained access control in cloud- based PHR systems. The original ABE system supports only monotone access policy and depends on a single private key generator. However, much work has been done in enhancing

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the ABE scheme. (Man HoAu, 2017) provide a general framework for secure sharing of PHR data. They modify the ABE scheme to encrypt access control information which allows stakeholders to encrypt and decrypt data without losing privacy preserving mechanisms. (Liu, 2018) present a modified Comparison based Encryption technique (CBE). The proposed hierarchical comparison-based encryption (HCBE) scheme incorporates an attribute hierarchy into CBE. The proposed scheme encrypts a ciphertext with a small number of generalized attributes at a higher level rather than many specific attributes at a lower level, greatly improving the encryption performance.

The work presented in this paper addresses the requirements for deploying a secure, low cost low power cluster for storage/ retrieval of personal healthcare records. We provide architecture of a hybrid cloud based PHR system that allows secure access and sharing of PHR based on access control parameters. The proposed security mechanism is based on Attribute based Encryption scheme. Our work is unique in terms of deployment of a power efficient system using a low-cost cloud computing cluster.

7. CONCLUSION AND FUTURE WORK

In this work, we have presented a design and architecture of a hybrid cloud environment for personal health records management using a low cost low power cluster built using two kinds of Single Board Computers. Details for requirements specifications for the proposed system were presented by classifying the requirements at three levels of data handling and data privacy: i) Data acquisition, ii) Data storage in Cloud, iii) Data Processing, analytics and visualization. Based on these requirements, a light weight architecture coupling the public and private cloud implementations was presented. Detailed aspects of encryption mechanisms based on ABE scheme were also presented. A low lost low power Hadoop cluster was deployed using two kind of single board computers, Raspberry Pi 3B and Odroid Xu4. Various experiments were conducted to evaluate the performance of the LoC4 cluster for computation time, storage management and network communications. Response time and task execution time were computed for various sizes of datasets to validate the security mechanism using a virtual machine in the proposed system.

Results from these studies show that a SBC based LoC4 cluster can be successfully deployed in a hybrid cloud environment. Although RPi computers are 50% cheaper than the Odroid Xu-4, the Odroid computer outperform RPi computers in all aspects of performance parameters. Xu-4 performs 3.8 times faster for Task execution times, 2.4 times faster for buffered read and write, 1.8 times faster for network throughput and at least 3 times faster for Hadoop based job executions. While SBC based clusters are energy efficient overall, the operation cost to performance ratio can vary based on the workload compared to a Personal Computer. In terms of power efficiency, for smaller workloads the Xu4 clustered devices performs 1.7 times better compared to RPi devices. Based on these observations, we conclude that LoC4 cluster is ideal for a hybrid cloud deployment where light weight applications can run on the back end. For heavier workload application, such as big data applications, due to the inefficient capabilities of these devices compared to high end server, the SBC based clusters may not be a performance effective choice. It is however possible to tweak Hadoop configuration parameters to adjust with given resources to improve the overall performance.

We provide response time evaluation experiments to validate the proposed scheme for encryption of data blocks based of ABE scheme. Results from the implementation of ABE based hybrid cloud scheme show on average a 43% improvement compared to AES based scheme for data block sizes of 1.8 MB to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

400 MB. The proposed encryption scheme performs 2.3 times faster compared to AES based scheme. The current implementation is limited to xml based data files mimicking a PHR data record. We intend to extend the proposed system to include Apache Hive based Database on the back-end LoC4 cluster with further deployment of MySQL based implementation in the virtual machine running in the public cloud. We also did not consider using secure communication channels. The current system relies on network sockets for listening and responding over TCP/IP protocol. In the future, we intend to incorporate secure and reliable communications mechanisms in the proposed system. We intend to further investigate the effects of various QoS parameters on the communication in the proposed system.

ACKNOWLEDGMENT

This work is partially funded by Research and Innovation Center at Prince Sultan University.

REFERENCES

Abbas, H., Maennel, O., & Assar, S. (2017). Security and privacy issues in cloud computing. *Annales des Télécommunications*, *72*(5-6), 233–235. doi:10.100712243-017-0578-3

Alghabban, W., Salama, R., & Altalhi, A. (2017). Mobile cloud computing: An effective multimodal interface tool for students with dyslexia. *Computers in Human Behavior*, *75*, 160–166. doi:10.1016/j.chb.2017.05.014

Apache Hadoop Framework. (2017). Retrieved from https://hadoop.apache.org/Hadoop

Asija, R., & Nallusamy, R. (2016). Healthcare SaaS Based on a Data Model with Built-In Security and Privacy. *International Journal of Cloud Applications and Computing*, *6*(3).

Au, M. H., Yuen, T. H., Liu, J. K., Susilo, W., Huang, X., Xiang, Y., & Jiang, Z. L. (2017). A general framework for secure sharing of personal health records in cloud system. *Journal of Computer and System Sciences*, *90*, 46–62. doi:10.1016/j.jcss.2017.03.002

Bahga, A. & Madisetti, V.K. (2015), Healthcare Data Integration and Informatics in the Cloud. *Computer*, *48*(2), 50-57.

Basu, S., Karp, A. H., Li, J., Pruyne, J., Rolia, J., Singhal, S., ... Swaminathan, R. (2012). Fusion: Managing Healthcare Records at Cloud Scale. *Computer*, *45*(11), 42–49. doi:10.1109/MC.2012.291

Baun, C. (2016). Mobile Clusters of single board computers: An option for providing resources to student projects and researchers. *SpringerPlus*, *5*(1), 360. doi:10.118640064-016-1981-3 PMID:27064532

Diffie, W., & Hellman, M. (2006). New directions in cryptography. *IEEE Transactions on Information Theory*, *22*(6), 644–654. doi:10.1109/TIT.1976.1055638

Fang, R., Pouyanfar, S., Yang, Y., Chen, S., & Iyengar, S. (2016). Computational Health Informatics in the Big Data Age: A Survey. *ACM Computing Surveys*, *49*(1). doi:10.1145/2932707

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gou, Z., Yamaguchi, S., & Gupta, B. B. (2016). Analysis of various security issues and challenges in cloud computing environment: a survey. In Handbook of research on modern cryptographic solutions for computer and cyber security (pp. 393-419). Hershey, PA: IGI Global. doi:10.4018/978-1-5225-0105-3.ch017

Health Insurance Portability and Accountability Act of 1996. (1996). Retrieved from https://www.hhs.gov/hipaa

Li, M., Yu, S., Zheng, Y., Ren, K., & Lou, W. (2013). Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *IEEE Transactions on Parallel and Distributed Systems*, *24*(1), 131–143. doi:10.1109/TPDS.2012.97

Liu, X., Liu, Q., Tao, P., & Wu, J. (2017). Dynamic access policy in cloud-based personal health record (PHR) systems. *Information Sciences*, *379*, 62–81. doi:10.1016/j.ins.2016.06.035

Nepal S., Ranjan R. & Choo K. (2015). Trustworthy Processing of Healthcare Big Data in Hybrid Clouds. *IEEE Cloud Computing*, *2*(2), 78-84.

Odroid Xu-4 detailed specifications (2017). Retrieved 1 September 2017 from http://www.hardkernel.com/main/products/prdt_info.php?g_code=G143452239825

Ostrovsky, R., Sahai, A., & Waters, B. (2007). Attribute-based encryption with non-monotonic access structures. In ACM conference on Computer and communications security (CCS '07) (pp. 195-203). doi:10.1145/1315245.1315270

Ostrovsky, R., Sahai, A., & Waters, B. (2007), Attribute-based encryption with non-monotonic access structures. In Proceedings of ACM CCS '07 (pp. 195-203).

Page, A., Kocabas, O., Soyata, T., Aktas, M., & Couderc, J. (2014). Cloud-based privacy-preserving remote ECG monitoring and surveillance. *Annals of Noninvasive Electrocardiology*, *20*(4), 328–337. doi:10.1111/anec.12204 PMID:25510621

Page, A., Soyata, T., & ... . (2015), Visualization of health monitoring data acquired from distributed sensors for multiple patients. In *Proc. IEEE Global Telecommun. Conf.* 10.1109/GLOCOM.2015.7417414

Peyton, L., Bindra, J., Baarah, A., Chamney, A., & Kuziemsky, C. (2015). A Bounded Health Information Technology System Design Approach to Support Community-Based Care Delivery. *International Journal of Cloud Applications and Computing*, *5*(1), 32–45.

Qureshi, B. (2014). Towards a Digital Ecosystem for Predictive Healthcare Analytics. In *Proceedings of the 6th International Conference on Management of Emergent Digital EcoSystems (MEDES '14)* (pp. 34-41). 10.1145/2668260.2668286

Raspberrypi.org. (n.d.). Raspberry Pi Single Board Computer Model 3B. Retrieved from https://www.raspberrypi.org/products/raspberry-pi-3-model-b

Soyata, T., Copeland, L., & Heinzelman, W. (2016). RF energy harvesting for embedded systems: A survey of tradeoffs and methodology. *IEEE Circuits and Systems Magazine*, *16*(1), 22–57. doi:10.1109/MCAS.2015.2510198

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sun, I., & Reddy, C. (2013). Big data analytics for healthcare. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '13)* (pp. 1525-1525). DOI: 10.1145/2487575.2506178

Venkatasubramanian, K., Banerjee, A., & Gupta, S. (2010). PSKA: Usable and secure key agreement scheme for body area networks. *IEEE Transactions on Information Technology in Biomedicine*, *14*(1), 60–68. doi:10.1109/TITB.2009.2037617 PMID:20007032

Wan, J., Zou, C., Ullah, S., Lai, C.-F., Zhou, M., & Wang, X. (2013). Cloud-enabled wireless body area networks for pervasive healthcare. *IEEE Network*, *27*(5), 56–61. doi:10.1109/MNET.2013.6616116

Wang, C., Xu, X., Shi, D., & Lin, W. (2014). An Efficient Cloud-Based Personal Health Records System Using Attribute-Based Encryption and Anonymous Multi-receiver Identity-Based Encryption. In *Proceedings of the Ninth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, Guangdong (pp. 74-81).

Wang, X., Gui, Q., Liu, B., Jin, Z., & Chen, Y. (2014). Enabling Smart Personalized Healthcare: A Hybrid Mobile-Cloud Approach for ECG Telemonitoring. *IEEE Journal of Biomedical and Health Informatics*, *18*(3), 739–745. doi:10.1109/JBHI.2013.2286157 PMID:24144678

Yi, L., Zhang, Y., Ling, J., & Liu, Z. (2018). Secure and fine-grained access control on e-healthcare records in mobile cloud computing. *Future Generation Computer Systems*, *78*(Part 3), 1020–1026.

Zhang, Y., & Lyu, M. R. (2017). *QoS Prediction in Cloud and Service Computing: Approaches and Applications*. Springer. doi:10.1007/978-981-10-5278-1

Zhou, J., Cao, Z., Dong, X., & Lin, X. (2015). PPDM: A Privacy-Preserving Protocol for Cloud-Assisted e-Healthcare Systems. *IEEE Journal of Selected Topics in Signal Processing*, *9*(7), 1332–1344. doi:10.1109/JSTSP.2015.2427113

Zhou, J., Lin, X., Dong, X., & Cao, Z. (2015). PSMPA: Patient Self-Controllable and Multi-Level Privacy-Preserving Cooperative Authentication in Distributed m-Healthcare Cloud Computing System. *IEEE Transactions on Parallel and Distributed Systems*, *26*(6), 1693–1703. doi:10.1109/TPDS.2014.2314119

ENDNOTES

1 https://wiki.gentoo.org/wiki/Sysbench
2 Using sysbench --test=cpu --cpu-max-prime=10000 --num-threads=*n* run
3 https://www.openhub.net/p/fio
4 https://iperf.fr/
5 http://bitspjoule.org/netpipe/

This research was previously published in the International Journal of Cloud Applications and Computing (IJCAC), 8(2); edited by B. B. Gupta and Dharma P. Agrawal; pages 27-46; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 34
EEMI – An Electronic Health Record for Pediatricians: Adoption Barriers, Services and Use in Mexico

Juan C. Lavariega
Escuela de Ingenieria y Tecnologias de Informacion, Mexico

Roberto Garza
Escuela de Ingenieria y Tecnologias de Informacion, Mexico

Lorena G Gómez
Escuela de Ingenieria y Tecnologias de Informacion, Mexico

Victor J. Lara-Diaz
Escuela de Ingenieria y Tecnologias de Informacion, Mexico

Manuel J. Silva-Cavazos
Escuela de Ingenieria y Tecnologias de Informacion, Mexico

ABSTRACT

The use of paper health records and handwritten prescriptions are prone to preset errors of misunderstanding instructions or interpretations that derive in affecting patients' health. Electronic Health Records (EHR) systems are useful tools that among other functions can assists physicians' tasks such as finding recommended medicines, their contraindications, and dosage for a given diagnosis, filling prescriptions and support data sharing with other systems. This paper presents EEMI, a Children EHR focused on assisting pediatricians in their daily office practice. EEMI functionality keeps the relationships among diagnosis, treatment, and medications. EEMI also calculates dosages and automatically creates prescriptions which can be personalized by the physician. The system also validates patient allergies. This paper also presents the current use of EHRs in Mexico, the Mexican Norm (NOM-024-SSA3-2010), standards for the development of electronic medical records and its relationships with other standards for data exchange and data representation in the health area.

DOI: 10.4018/978-1-7998-1204-3.ch034

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

Mexico, like other developing countries, holds a social debt for its citizens with respect to the provision of health services. In particular, one of the most vulnerable groups is the infant population. Even though the global child mortality rate (considering children under five years old) has been declining from 90 to 48 deaths per 1000 live births between 1990 to 2012 (WHO 2013), there is still work to be done. In urban parts of Mexico, the infant mortality rate is 16.2 deaths per 1000 live births, but the number in rural areas is even higher, where children present symptoms of malnutrition, untracked growth and untracked immunizations. However, more than a half of child deaths in general (rural and urban areas) are due to diseases that are preventable and treatable through simple and affordable interventions. Some of the deadliest childhood diseases like measles, polio, diphtheria, tetanus, pertussis and pneumonia have immunizations available that can protect children from illness and eventually, death (WHO 2013).

Electronic Health Records (EHRs) systems with key functionality for pediatricians can help reduce these problems (malnutrition, untracked growth and untracked immunizations) by recording a child's information, generating immunization schedules and comparing child growth with recommended world charts (WHO 2007). EHRs assist pediatricians, who, as many other medical professionals, have a heavy work load and a priority for up to date knowledge. Pediatricians manage the physical, behavioral and mental health of children from birth up to age 21. They are trained to diagnose and treat a broad range of childhood illness, ranging from minor health problems to serious diseases. Unfortunately, pediatricians just like any human being, are susceptible to make mistakes, can feel tired and/or be distracted, and then write incorrect dosages, misspell a medication name or simply write an illegible prescription (Bates et al 2001).

The National Coordinating Council for Medication Error Reporting and Prevention (NCCMERP) defined a medication error as "any preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of the health care professional, patient, or consumer (NCCMERP 2015). According to the American Society of Health-System, Pharmacists prescription errors are any incorrect drug selection, dose, dosage form, quantity, concentration, rate of administration, instructions of use, and any illegible prescriptions that lead to errors (ASHP 1993). The dose selection errors represent more than 50% of all prescribing faults, but other possible errors could be due to inaccuracy in writing, poor legibility of handwriting, use of abbreviations or incomplete writing of a prescription because all of these errors can lead to misinterpretation by pharmacists and patients (Velo, G.P., & Minuz, P. 2009).

In a study conducted at a Mexican university (Zavaleta-Bustos, M., et al., 2008), a sample of 370 prescriptions were randomly selected for analysis. The results indicated that 214 prescriptions (58%) of the sample had at least one error. The most common errors were incorrect indication, patient allergy to a medication, incorrect dosification, unjustified medications and duplicity of medications. Many of these errors were produced during the act of writing out a prescription.

Pediatricians, when confronted with an adult-oriented EHR, complain about not having specific functions for child medical care. The pediatricians' software requirements for EHRs are: immunization management, growth tracking, and medication dosing. Without these functions in an EHR, pediatricians are not capable to provide with quality care (Spooner, S. A., et al 2007).

An EHR system that fully supports pediatric practice must allow the recording of multiple immunizations and be able to analyze immunization data to determine the age when each vaccine should be administered. In addition, the system must provide graphic information of a child's body measurements

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(weight, height, head circumference, body mass index) over time, as pediatricians make judgments based on these measurements. Another requirement is the method for calculating drug dosages; the system must be able to calculate doses based on weight or age depending of the medication (Spooner, S. A., et al 2007). Based on these needs, EHR should be adopted as a helpful tool for physicians. However, in developed countries existing EHRs are expensive, difficult to use, developed from the hospital administration point of view, or do not meet physician's needs. As a consequence, unless there is a clear benefit (economical or better job status) for using EHR, physicians do not use EHRs and usually see these systems as a waste of time.

This papers presents EEMI (abbreviated by its name in Spanish, Expediente Electrónico Médico Infantil or Child Electronic Health Record) a system for the administration of medical records planned for pediatric medical practice. EEMI was designed and tested with the active participation of experienced pediatricians working at one of the top hospitals in the metropolitan area of Monterrey, Mexico. The objective of EEMI is to help pediatricians to minimize the possibility of errors, optimize consultation time (spend more time with patient, rather than filling forms) and keep track of patients' information.

The rest of this paper is organized as follows. Section 2 presents the use of EHR in context, including the barriers that this type of system faces during its adoption or rejection by medical professionals; a brief discussion of standards and international norms for EHRs; and a comparison of current EHRs including EEMI. Section 3, formally presents EEMI through the description of its functionality. Section 4 describes the technology used in EEMI. Finally, Section 5 discusses current status and future work.

2. BACKGROUND AND RELATED WORK

This section includes a description of the adoption barriers that EHRs face with the physician community; describes the standards and regulations currently available for EHR's; and concludes with a brief description of the main EHRs used in the US and in Mexico.

2.1. Adoption Barriers

According to some studies (Humpage, S. 2010, NCCMERP 2015, Spooner, S. A., et al 2007, Velo, G.P., & Minuz, P. 2009), and our own experience (Quintas Ruiz I, 2011) the use of EHRs, increases productivity by 20%, reduces waiting time up to 60% and saves 80% of paper work cost. Additionally, EHRs have other benefits such as: increasing security on patient sensitive data; providing easy and fast access to patient information; decreasing negative medical events (overmedication, incorrect treatment, etc.); and decreasing cost by unnecessary or repetitive treatments or laboratory tests. Despite their benefits, EHRs have not been fully adopted due to factors such as change resistance, cost, lack of incentives and complex customization (Hsiao C.J., & Hing E.2013, Humpage, S. 2010, Spooner, S. A., et al 2007, Spooner, S. A., 2007, Quintas Ruiz I, 2011).

2.1.1. Diverse Technology and Adoption Speed

There is no clear measure of how successful is the use of EHR, because doctors, patients and hospitals are using diverse technologies, and some of them are not using any at all. Some EHR systems run on standalone machines, others use mobile devices and internet services.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.1.2. Resistance to Change

Resistance to change is the main barrier for adoption, particularly by physicians. Arguments given by physicians include system cost, job overload and concerns about information security and privacy. Physicians which use EHR, do because they have been instructed by their superiors or see a benefit, for example an economic bonus or a job promotion.

2.1.3. Cost

Cost is another barrier for adoption of EHRs. Most of electronic health record systems are implemented in large medical centers with complex functionality; from administrative activities management (insurance processing, billing, patient's registration, etc.) to health information management (patient medical history, lab analysis, surgery schedules, etc.). Therefore, the cost of EHR increases and only large medical centers can afford to have electronic health records. It is difficult to quantify the benefits that Physicians find in their practice when using EHR. It is hard to evaluate the increase in productivity, cost reduction, and the return of investment (ROI).

2.1.4. Incentives Wrongly Aligned

Cost savings by using an EHR may not have a direct impact in the physician income, but in the patient cost or in the health insurance company earnings. For example, an EHR can alert when a medicine is prescribed to an allergic person, then the patient or his/her insurance company can save money with the avoidance of the need for further anti-allergy-related treatment.

2.1.5. Personalization and Work Flow

Successful systems are customizable to different types of users (physicians in this case) but such a customization may be complex and time consuming. Additionally, physicians have a personal style to do their job and they see EHR as an unwelcome change in the way they work, affecting their productivity.

During the design and development of EEMI, pediatricians worked directly with the development team in order to correctly map their needs into EEMI functionality, minimize resistance to change and technology adoption. Deployment of EEMI was done in increments; doctors had a functional version of the system, on every release. Also workflow templates were generated for a less complex customization and a successful implementation.

2.2. Health Standards and the Mexican Norm

EHRs are nowadays subject to regulations with respect to how they represent, store and share information with other systems. A brief description of the Mexican norm for electronic health systems and international health information standards follows.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.2.1. Norma Oficial Mexicana

The Mexican Norm (Norma Oficial Mexicana NOM-024-SSA3-2010 and NOM-024-SSA3-2012) (Secretaría de Salud. 2010, 2011) are the standards that set the objectives and functional requirements that an electronic health record must have in order to guarantee the interoperability, confidentiality, security, and information catalogs. Among its specification the Mexican Norm states that electronic health records have to demonstrate that a minimal set of data is generated. The Mexican Norm requires the use of international standards for disease classification (ICD), interoperability (HL7) and medical images format (DICOM). Systems that are regulated by the norm include those used in outpatient consultation, inpatient care, emergency care, pharmacy, laboratories, surgery, and imaging.

2.2.2. International Classification \of Diseases (ICD)

The International Classification of Diseases (ICD) (WHO 2015) is the standard diagnostic tool for epidemiology, health management and clinical purposes. This includes the analysis of the general health situation of population groups. It is used to monitor the incidence and prevalence of diseases and other health problems, proving a picture of the general health situation of countries and populations. The ICD is published by the World Health Organization. The Mexican Norm requires that EHR use a disease catalog such as the ICD.

2.2.3. Logical Observation Identifiers Names and Codes (LOINC)

Logical Observation Identifiers Names and Codes (LOINC) (LOINC 2015) is the standard for identifying medical laboratory observations. LOINC includes categories for chemistry, hematology, and microbiology, among others in its laboratory section. The clinical observations section includes vital signs, electrocardiograms and echocardiograms among other categories.

2.2.4. Digital Imaging and Communications in Medicine (DICOM)

Digital Imaging and Communications in Medicine (DICOM) (DICOM 2015) is the standard for medical images. DICOM defines how to handle, store, print and transmit information in medical imaging. It includes a file format definition and a network communications protocol. This standard is implemented in almost all devices used in radiology, cardiology, and ultrasound. The standard is also being applied to ophthalmology and odontology. The Mexican norm, mentions DICOM as the standard to be used for medical images by EHRs.

2.2.5. Health Level Seven (HL7)

Health Level Seven (HL7) (Benson, T 2010, Health Level Seven International 2015) is a set of standards for the transference of clinical and administrative data between different health software applications including EHRs. HL7 refers to the focus of the standards in the application layer (or layer 7) in the Open System Interconnection (OSI) model. There are two versions: HL7 Version 2 and HL7 Version 3 from which Version 2 is the most widely used in health related systems. Communication between applications as defined in HL7 is through well-formed text documents. Mexican Norm requires data exchange between health systems to follow the HL7 standard.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.3. EHR Systems Comparison

In the US, the adoption of EHR has increased in the last years, in part due to government incentives, maturity of the EHRs, and better acceptance by medical professionals. EHR enterprises have perceived the market growth and therefore they have developed a great variety of products, as shown in Table 1 (EPIC 2015, Cerner 2015, Allscripts 2015, DrChrono Inc. 2015, Entrada 2015, Smart EMR 2015, eClinicalWorks 2015, PracticeFusion Inc 2015, Athenahealth, Inc 2015, iCare 2015, Bush, J. & Baker, S 2014, Atlassian.2015).

In Mexico, after the release of the Mexican Norm for Electronic Health Records in 2010) (Secretaría de Salud. 2010, 2011), new EHR's have been developed. The majority of these EHR implementations are in public health system hospitals and some in private medical centers. Public institutions such as IMSS (Social Security Institute for Mexico), ISSSTE (Social Security Institute for State Workers), PEMEX Hospitals (Mexican National Oil Company), Secretaría de Marina (Mexican Navy) and the private ABC Hospital have their in-site developed EHR. There are few commercial systems such as eMedix, Med2k, and, Alert. None of the reviewed systems for Mexico has a patient version or portal. Table 2 shows the characteristics of EHR in Mexico (ISSSTE. 2010, Estado de Colima 2015, Secretaría de Marina Armada de México 2015, Yacamán, R. 2010, Camacho Ortega V. 2014, eMedix 2015, Med2k 2015, Alert 2015).

Table 1. EHRs in the USA

EHR	Deployment	Focus on	e-prescribing	Standards	Comments
EPIC	On Premise	Hospital	Prescription creation	HL7	Leader in the Gartner 2013 Magic quadrant for Enterprise EHR systems.
Cerner	Web, Mobile	Patient	NA	NA	Leader in the Gartner 2013 Magic quadrant for Enterprise EHR systems.Personal Health Record. Shows metrics in a graph.
AllScripts	Web	Hospital	Pharmacy link	NA	Visionary in the Gartner 2013 Magic quadrant for Enterprise EHR systems.
DrChrono	Web, Mobile	Physician	Prescription creation, Pharmacy link	ICD-9	Patients can access their demographic information, manage appointments, query lab results, send and receive messages to physicians, pre-fill forms in advance. Shows metrics in a graph.
SmartEMR	Web	Physician	Prescription creation, link to Surescripts (largest e-prescribing network)	NA	Patient can review their medical history, receive appointment reminders, fill information in advance, request appointments, review lab results, request medicine refills and communicate with physicians. Shows metrics in a graph.
eClinicalWorks	Web, Mobile	Hospital	Prescription creation, Pharmacy link	ICD-9	Patients can manage appointments, query their medical record, access lab results, request medicine refills and access educational information.
Practice Fusion	Web, Mobile	Physician	Prescription creation, Pharmacy link	ICD-9	Patients can manage appointments, query lab results, consult prescriptions, consult immunization schedule and send messages to physician. Shows metrics in a graph.
Athena Health	Web, Mobile	Physician	NA	ICD-9	Patients can manage appointments, receive appointment reminders, and consult lab results, Shows metrics in a graph.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. EHRs in Mexico

EHR	Deployment	Focus on	e-prescribing	Standards	Comments
IMSS	On Premise	Hospital	Prescription creation	NA	This system is only used for patients who are affiliated with IMSS.
eMedix	Web	Physician	NA	ICD-9, ICD-10	
ISSSTEMed	On Premise	Hospital	Prescription creation	NA	Interaction with pharmacy system and lab test, This system is only used for patients who are affiliated with ISSSTEMed.
SAECCOL	On Premise	Hospital	NA	ICD-9, ICD-10	This system is used by public hospitals in the state of Colima
SICOHOSP	On Premise	Hospital	Prescription creation	ICD-10	Interaction with lab test system
Hospital ABC	On Premise	Hospital	NA	NA	
PEMEX	On Premise	Hospital	ICD-10	ICD-10	
Alert	On Premise	Hospital	Prescription creation	HL7	
Med2K	On Premise	Physician		ICD-10	Appointment administration, prescription, image storage, video and voice.
EEMI	Web	Physician	Prescription creation	ICD-10	Automatic creation of immunization schedule.

The information summarized in Table 1 and Table 2 helped us to define and design the functionality on EEMI. One of the key factors we found in the analysis was the need to keep record of immunizations, create prescriptions automatically but in an editable by physicians' format and of course the imperative to follow the Mexican Norms (Secretaría de Salud. 2010, 2011) as well as international standards. In the following sections, a detailed discussion on EEMI is presented.

3. EEMI SYSTEM DESCRIPTION

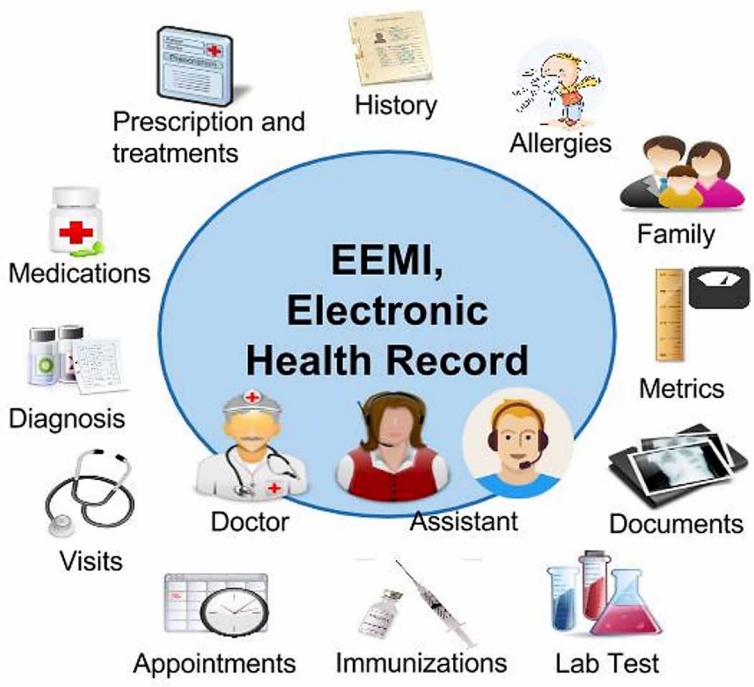
EEMI is an electronic health record system based on the practice experience of pediatricians in Monterrey, Mexico. The system follows the Mexican norm in its functionality and data exchange. Figure 1 represents a contextual overview of the services provided by EEMI. Users of the Child Electronic Health Record, presented in this paper, are physicians and their office assistants. The following paragraphs describe EEMI services.

3.1. Appointment Administration

This module allows to creates and manage appointments. This service resolves date conflicts with national holidays and personal physician calendar. The Physician can manage his/her own calendar to mark his/her vacations or define days out of the office.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. EEMI contextual overview



3.2. Visits Management

This module includes access to previous diagnosis, known treatments and related medications to facilitate and speed up the physician consultation. The system is able to automatically generate a prescription based on the diagnostic and calculate the medication dosage. The automatically generated prescription can be edited by the physician according to his/her consideration.

3.3. Medical Record Management

This module allows the management of information such as physical measurements, medical history, family, documents, laboratory results, immunizations, appointments, and visits history. New information can be added or current information can be updated.

3.4. Diagnosis

This module allows the specification of visit diagnosis based on the International Classification of Diseases (WHO 2015)

The system has a section for the diagnosis, treatments, medications, medications generic and brand names. This information is presented in form of lists that are related with each other and whenever a physician selects a diagnosis the system will display the possible treatments for that diagnosis, when a physician selects a treatment the system will display the possible medications for that treatment and the same goes for the medications generic and brand names.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Every time a physician selects a diagnosis the system will learn the relationships between diagnosis, treatment, medication and brand names and also the information for each one of them. This eliminates the need for the physician to type the same information every time for the same diagnosis. However, the possibility of modifying the selected treatment is kept always open, since there are no fixed combinations of diagnostic-treatment-medication-generic-brand name.

Each physician can create, modify and save his/her own catalogs of diagnosis, treatments, medications generic and brand names and the information could be used by another physician but could also be modified in case of an error, or a different treatment preference, choice of medication or a different brand name. The creation of this medical collaboration will reduce the errors in diagnosis.

Additionally, the system identifies frequent diagnosis in the last "n" days and every time a physician selects one of them, the system will suggest a full diagnosis, set of diagnosis-treatment-medication and brand name based on the most recently used. This suggestion makes faster the process of consulting a patient and is also helpful because many diseases have different prevalence depending on the season of the year and have a tendency to appear as clusters.

Every time a medication is added to the prescription, the system validates that the medication doesn't have an incompatibility with other medications in the same prescription. EEMI also validates medication against patient's known allergies.

3.5. Prescriptions and Dose Calculation

Once the diagnosis is selected, the system creates the prescription based on the information of every treatment and medication; it calculates the dose and generates a text for each medication that contains the presentation, dose, administration route, dispensing time and duration of treatment. The physician can print the prescription and give the patient a legible prescription that will not cause any misinterpretation. The dose calculation for every medication can be calculated by body weight or patient's age.

3.6. Immunization Schedule

The immunization schedule is automatically programmed based on the recommendations by the Mexican Ministry of Health (SSA). The doctor can personalize this schedule adding new immunizations not considered by the SSA. The system records the immunization and the age range of application. The recommended immunization date will be used for calculating the estimated date of application for a child and will create an appointment for that date. When a child is born and visits the pediatrician for the first time, the physician's assistant uses the system to schedule a full calendar of immunizations and appointments. This automation helps to keep track of the child's immunizations and allows the assistant and the physicician to remind the family for future visits. EEMI sends automatic e-mail remainders to parents or tutors of child's immunizations appointments.

3.7. Growth Graphs

The system displays the child's anthropometric measurements (weight, height, head circumference, body mass index) in a percentile chart, so the physician can assess the growth over time but is also able to compare with statistical data, because the system also displays the percentile distribution of all population values.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

With EEMI, the pediatrician can view, print or save in various graphic formats data of weight for age, height for age, weight for height, head circumference for age, and body mass index for age.

In a weight for age graph the pediatrician can easily identify when a child presents overweight or malnutrition by just comparing the child weight and check if the value is outside the normal values. Figure 2 shows an example of a low weight for age graph, where the blue dots are the child's weight and the color lines are the normal values.

3.8. EEMI Benefits

The following are the benefits of EEMI adoption according to observations by the initial testing group. Most of the benefits are not directly quantifiable in monetary terms, but qualitative in terms of a better interaction between physicians and patients.

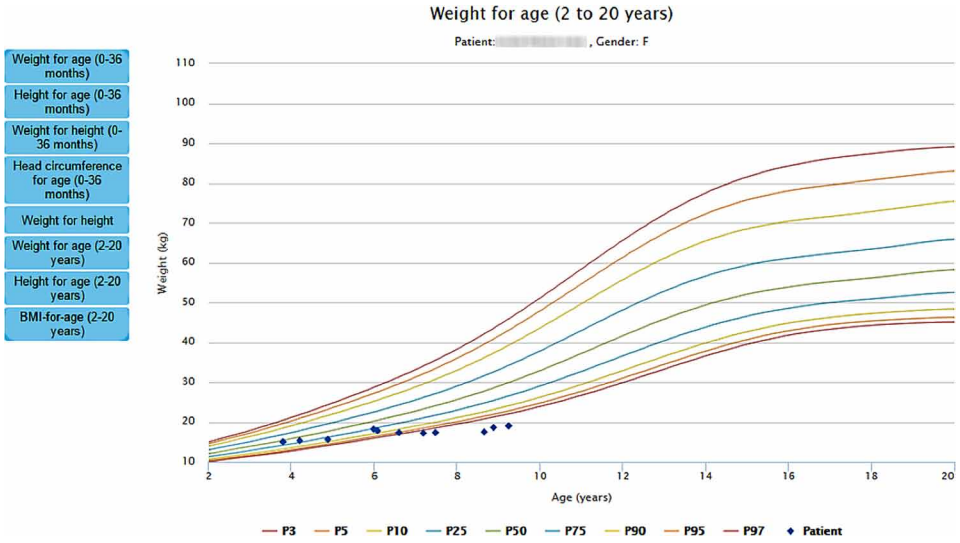
3.8.1. Tax Deductible Cost

In its current implementation as a service, EEMI is tax deductible by the Mexican fiscal legislation. Therefore, there is no a direct economic cost to the physicians that are using EEMI. Other EHR that were developed as standalone desktop applications are not considered services by the fiscal law, and the investment and other monetary cost are covered by the adopters

3.8.2. Optimization of Visit Time

During the visit, physicians spend more time talking with patients and understanding his/hers needs that looking for past visits records or writing down prescriptions and indications.

Figure 2. Growth graph (weight vs age)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.8.3. Elimination of Prescription Errors

One key functionality of EEMI is the automatic (and editable) creation of prescriptions. This feature reduces common errors that physicians may do while handwriting or even typing prescription. It also allows to reduce or even eliminate medication incompatibility errors by looking for contra-indications for medicines in the prescription.

3.8.4. Patient Data Consistency and Privacy

Other key benefit mentioned by adopters of EEMI is the consistency in the information that is included in the prescriptions, such as medicines, dosage and particular indications. By having all this information at hand from visit to visit allows physicians a better understanding of the patients' progress.

3.8.5. Elimination of Paper Storage

The use of EEMI has permitted to reduce or eliminate the physical space required for storage of paper records.

3.9. EEMI Adoption and Barriers

EEMI has been evaluated by a group of physicians during its testing period. Even though, the acceptance of EEMI has been generally good, there are still some adoption barriers among the medical community. The most prevalent are:

Change resistance. As in any professional group, there is a group of physicians that is not willing to accept an EHR of any kind, and keeps attached to the traditional way to do a consult.

Use customization. One aspect for a good acceptance of an EHR according to the evaluation group has been the customization of the systems to the physician practice. EEMI has been successfully adapted to the way pediatricians work, however it is possible that other physicians find EEMI workflow inadequate to their practice.

Information transfer. Another key element for delay in the adoption of EEMI has been the burden of the initial transfer of patient information from one legacy system to EEMI. Without this feature, physicians face the dilemma of capturing (typing) again all their patient records or to keep the legacy systems for old time patients and record the new patients in EEMI,

4. EEMI TECHNOLOGY DESCRIPTION

EEMI is built based on the N-Tier Architecture, specifically 3-Tier Architecture. It's a client-server architecture where the logic, view and data are in different tiers. The data tier can contain one or more databases, the business tier contains all the logic of the project and the presentation tier contains the user interface. Currently EEMI is being implemented on a Service Oriented Architecture (SOA) (Schuldt, H. 2009) that provides a loosely coupled approach that allows interoperability, maintenance and updates. The use of Web Services allows data sharing between applications.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A study showed that physicians have an urgent requirement to access the electronic health record systems with personal devices such as tablets or smartphones; therefore, a Service Oriented Architecture is suggested in the study (Wu, Y., & Yang, H. 2012). With this architecture we can develop native applications for diverse operating systems (iOS and Android) that cover the mobile devices alternatives in the market. This is possible by only developing the presentation tier and using the Web services.

The Figure 3 shows the architecture of the system, which is composed of three tiers. The presentation tier contains all the interfaces that presents data to the end users and allows the data manipulation. The business logic tier contains all the web services that will obtain and modify the database. The data tier is responsible for the database management. This separation of tiers improves scalability, reusability, flexibility, and maintainability, because we have loosely coupled and highly cohesive methods in each tier that are easy to change or reuse. Is important that the system has these attributes because we want to extend this system to other medical areas like gynecology.

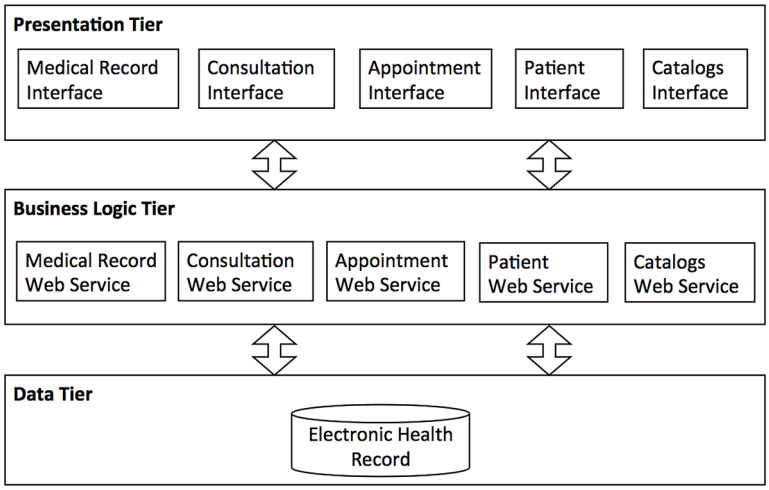
EEMI is implemented using .NET Framework 4.5 for the development of this system. The presentation tier was developed in ASP.NET, the business logic tier in Windows Communication Foundation (WCF) and for the data tier we use SQL Server 2012.

5. CURRENT STATUS AND FUTURE WORK

At the time of writing this paper, EEMI has been in the acceptance testing phase with a pilot in a group of pediatricians in Monterrey, Mexico. We will review the results and observations of this phase for include improvements in the production version. The production version was completed by the end of August 2015.

As future work, EEMI will be integrated with the project for Monitoring and Assisting Maternity-Infant Care in rural areas (MAMICare) (Lavariega, J., Córdova, G., Gomez-Martinez, G, & Ávila, A.2013). The MAMICare project monitors pregnant women to detect anomalies and prevent maternity and infant deaths. The EEMI system will continue with the monitoring of the child by keeping track of the growth and immunizations of the child.

Figure 3. SOA for EEMI



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

New electronic devices are being reviewed. These devices send information via Bluetooth to automatically register the measurements in the system. One key idea is to incorporate these devices in the process of taking measurements to eliminate the possibility of the injection or errors.

This system can be accessed from any tablet or smartphone because it is a Web application, but the interface may look different. One objective for the short terms is to develop native iOS and Android applications of this electronic medical record, a decision has to be made upon including only the most important requirements for these new applications or include everything that is available in the main EEMI system.

6. CONCLUSION

In general, EEMI enhances the process of a pediatric medical consultation due to the automation of several common activities. It also reduces the possibility of error that could be incurred any time the physician types a repetitive diagnosis. Due to the automated functions, the physician can focus on giving a better medical care. As a consequence of fewer typing errors, patient information is more accurate and a sense of security can be established between doctor and patients.

EEMI is a useful tool for keeping track of the child's growth and creating an immunization schedule for each child. The graphs will help the pediatrician to make decisions based on the weight, height, head circumference, and body mass index of a child.

With the use of a Service Oriented Architecture the system is scalable, reusable, flexible, and maintainable. This architecture makes it easier to make changes in the system, and reuse code because of the highly cohesive and loosely coupled methods. We can create native applications for iOS and Android by just developing the presentation tier and use the Web Services that are already developed in the business logic tier.

REFERENCES

Alert. (2015). Alert Online. Retrieved from https://www.alert-online.com/es/company

Allscripts. (2015). Professional EHR. Retrieved from www.allscripts.com

ASHP. (1993). American Society of Hospital Pharmacists ASHP Guidelines on Preventing Medication Errors in Hospitals. *American Journal of Hospital Pharmacy*, *50*(2), 305–314. PMID:8480790

Athenahealth, Inc. (2015). athenaClinicals. Retrieved from www.athenahealth.com

Atlassian.(2015) Jira. Retrieved from https://www.atlassian.com/software/jira

Bates, D. W., Cohen, M., Leape, L. L., Overhage, J. M., Shabot, M. M., & Sheridan, T. (2001). Reducing the frequency of errors in medicine using information technology. *Journal of the American Medical Informatics Association*, *8*(4), 299–308. doi:10.1136/jamia.2001.0080299 PMID:11418536

Benson, T. (2010). *Principles of Health Interoperability HL7 and SNOMED*. Health Informatics Series. doi:10.1007/978-1-84882-803-2

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bush, J., & Baker, S. (2014). Where Does It Hurt? An Entrepreneur's Guide to Fixing Health Care (1st ed.)

Camacho Ortega, V. (2014). El Expediente Clínico Electrónico en PEMEX: un caso de éxito. *Proceedings of the 4a Reunión Nacional de Información en Salud*. Retrieved from www.dgis.salud.gob.mx

Cerner. (2015). Cerner Personal Health Record. Retrieved from https://cernerhealth.com/

de Onis, M. (2007). Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization*, *85*(09), 660–667. doi:10.2471/BLT.07.043497 PMID:18026621

DICOM. (2015). About DICOM. Retrieved from http://medical.nema.org

Dr Chrono Inc. (2015). Dr Chrono. Retrieved from www.drchrono.com

eClinicalWorks (2015). eClinicalWorks v10 (EHR Suite). Retrieved from www.eclinicalworks.com

eMedix. (2015). eMedix Expediente Clínico Electrónico. Retrieved from www.emedix.com.mx

Entrada. (2015), Entrada EHR. Retrieved from www.entradahealth.com

EPIC. (2015). EpicCare Ambulatory Electronic Medical Record (EMR). Retrieved from www.epic.com

Estado de Colima. (2015). Saeccol sistema para la administración del expediente clínico electrónico colima. Retrieved from www.salud.gob.mx

Health Level Seven International. (2015). About HL7. Retrieved from www.hl7.org

Hsiao C.J., & Hing E. (2013). Use and characteristics of electronic health record systems among office-based physician practices: United States, 2001–2013. NCHS data brief, no 143.

Humpage, S. (2010). Benefits and Costs of Electronic Medical Records: The Experience of Mexico's Social Security Institute. Retrieved from http://www.iadb.org

iCare. (2015). The Enterprise Cloud EHR. Retrieved from www.icare.com

ISSSTE. (2010). *ISSSTEMed. Avances en la Implementación del Expediente Clinico Electronico.* Retrieved from www.salud.gob.mx

Lavariega, J., Córdova, G., Gomez-Martinez, G., & Ávila, A. (2013). Monitoring and Assisting Maternity-Infant Care in Rural Areas (MAMIcare). *Transactions of the International Conference on Health Information Technology Advancement*, 2(1), 175-184.

LOINC. (2015). History, Purpose, and Scope. Retrieved from www.loinc.org

Med2k. (2015). Med2k Expediente Electrónico. Retrieved from http://www.complise.mx/software.asp?producto=med2k

NCCMERP. (2015). National Coordinating Council for Medication Error Reporting and Prevention, About Medication Error. Retrieved from http://www.nccmerp.org/aboutmederrors.htm

PracticeFusion Inc. (2015). PracticeFusion EHR. Retrieved from www.practicefusion.com

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Quintas Ruiz, I. (2011). Analisis de Factores Críticos de Éxito para implantación de Expediente Clínico Electronico en el Área Metropolitana de Monterrey [Master of Administration on IT Thesis]. ITESM-Monterrey Campus.

Schuldt, H. (2009). Multi-Tier Architecture. Encyclopedia of Database Systems.

Secretaría de Marina Armada de México. (2015). *SISTEMA DE CONTROL HOSPITALARIO (SICOHOSP)*. Retrieved from www.salud.gob.mx

Secretaría de Salud. (2010). Norma Oficial Mexicana NOM024SSA32010. *Diario Oficial de la Federación. México DF: Secretaría de Salud*. Retrieved from www.dgis.salud.gob.mx

Secretaría de Salud. (2011). *Manual del Expediente Clinico Eletronico V2011*. Retrieved from http://www.who.int/goe/policies/countries/mex_ehealth.pdf

Smart EMR. (2015). smart EMR. Retrieved from www.smartemr.com

Spooner, S. A. (2007). Special requirements of electronic health record systems in pediatrics. *Pediatrics*, *119*(3), 631–637. doi:10.1542/peds.2006-3527 PMID:17332220

Spooner, S. A., & ... . (2007). Special requirements of electronic health record systems in pediatrics. *Pediatrics*, *119*(3), 631–637. doi:10.1542/peds.2006-3527 PMID:17332220

Velo, G. P., & Minuz, P. (2009). Medication errors: Prescribing faults and prescription errors. *British Journal of Clinical Pharmacology*, *67*(6), 624–628. doi:10.1111/j.1365-2125.2009.03425.x PMID:19594530

WHO. (2013). World Health Organization. *Children: reducing mortality updated at 2013*. Retrieved from http://www.who.int/mediacentre/factsheets/fs178/en/

WHO. (2015). International Classification of Diseases (ICD) *World Health Organization (WHO).* Retrieved from http://www.who.int/classifications/icd/en/

Wu, Y., & Yang, H. (2012). An electronic medical records review system for mobile healthcare based on web services. *Paper presented at the 5th International Conference on Biomedical Engineering and Informatics* (pp. 1040-1044). 10.1109/BMEI.2012.6513060

Yacamán, R. (2010). Expediente Médico Electrónico en el Centro Médico ABC. *3era Reunión Nacional de Información en Salud.* Retrieved from www.dgis.salud.gob.mx

Zavaleta-Bustos, M., & ... . (2008). Prescription errors in a primary care university unit: Urgency of pharmaceutical care in Mexico. *Brazilian Journal of Pharmaceutical Sciences*, *44*(1), 115–125.

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 11(3); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 57-69; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 35
Constructing a Bioinformatics Platform with Web and Mobile Services Based on NVIDIA Jetson TK1

Chun-Yuan Lin
Chang Gung University, Taiwan

Chung-Hung Wang
Chang Gung University, Taiwan

Jin Ye
Chang Gung University, Taiwan

Min Su
Beijing University of Technology, China

Che-Lun Hung
Providence University, Taiwan

Jianjun Tan
Beijing University of Technology, China

ABSTRACT

Current high-end graphics processing units (abbreviate to GPUs), such as NVIDIA Tesla, Fermi, Kepler series cards which contain up to thousand cores per-chip, are widely used in the high performance computing fields. These GPU cards (called desktop GPUs) should be installed in personal computers/servers with desktop CPUs; moreover, the cost and power consumption of constructing a high performance computing platform with these desktop CPUs and GPUs are high. NVIDIA releases Tegra K1, called Jetson TK1, which contains 4 ARM Cortex-A15 CPUs and 192 CUDA cores (Kepler GPU) and is an embedded board with low cost, low power consumption and high applicability advantages for embedded applications. NVIDIA Jetson TK1 becomes a new research direction. Hence, in this paper, a bioinformatics platform was constructed based on NVIDIA Jetson TK1. ClustalWtk and MCCtk tools for sequence alignment and compound comparison were designed on this platform, respectively. Moreover, the web and mobile services for these two tools with user friendly interfaces also were provided. The experimental results showed that the cost-performance ratio by NVIDIA Jetson TK1 is higher than that by Intel XEON E5-2650 CPU and NVIDIA Tesla K20m GPU card.

DOI: 10.4018/978-1-7998-1204-3.ch035

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

Current high-end graphics processing units (abbreviate to GPUs) become very popular in the high performance computing community fields. These GPU cards, such as NVIDIA Tesla, Fermi, Kepler series, contain up to thousand cores per chip. For example, NVIDIA Tesla K20m GPU card has 2496 CUDA cores. They contain massive multi-threaded processors; moreover, the thousands of threads can be declared and executed simultaneously to fully utilize GPU computing power. However, these GPU cards, called desktop GPUs, should be installed in a personal computer or a server with desktop CPUs. Moreover, the cost and power consumption of constructing a high performance computing platform with these desktop GPU cards are high. For example, a NVIDIA Tesla or Fermi GPU card may be needed to spend thousands of US dollars and a personal computer or a server may also spend about hundreds to thousands of US dollars. The overall power consumption by this platform may be up to a kilowatt. Besides, this platform is hard to provide the immediate and mobility requirements, i.e., the developers should demo the program by the remote control.

Jetson TK1 (or called Tegra K1, http://elinux.org/Jetson_TK1) released by NVIDIA is a full-featured platform for embedded applications and it contains 192 CUDA Cores (Kepler GPU), 4 ARM Cortex-A15 CPUs and 2 ISP cores. NVIDIA Jetson TK1 also contains Bluetooth and Wi-Fi modules for mobile applications. The cost of a NVIDIA Jetson TK1 is less than three hundred US dollars and its power consumption is about tens of watts by the tests (Paolucci et al., 2014). Several reports or literatures have been proposed to compare NVIDIA Jetson TK1 with other embedded platforms, desktop CPUs and desktop GPUs, respectively. In the NVIDIA Jetson TK1 Development Kit (http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html), it shows that the speedup ratios and the performance per watt by NVIDIA Jetson TK1 are better than those by Apple A7 used in the iphone5. James Wolfer (Wolfer, 2015) compared NVIDIA Jetson TK1 with Raspberry PI Model in terms of time and speedup ratios. The results showed that the single NVIDIA Jetson TK1 outperforms Raspberry PI Model in these two items. P.S. Paolucci *et al*. (Paolucci et al., 2014) also compared the ARM Cortex-A15 CPU in NVIDIA Jetson TK1 with the Intel XEON E5620 CPU in the SuperMicro server in terms of joule per synaptic event, total energy, power consumption, and execution time. For the first two items, NVIDIA Jetson TK1 achieves 4.5 and 4.4 times performance better than the SuperMicro server, respectively. NVIDIA Jetson TK1 achieves 14.4 times performance better than SuperMicro server for the power consumption. However, the execution time by NVIDIA Jetson TK1 is 3.3 times larger than that by the SuperMicro server. S. Fu *et al*. (Fu et al., 2015) do the comparison of performance, power efficiency, and cost efficiency among desktop Intel i7-3770 multi-core CPU, desktop NVIDIA GTX 690 GPU card and NVIDIA Jetson TK1. The results proved that the power efficiency and cost efficiency by NVIDIA Jetson TK1 are both better than those by desktop Intel i7-3770 CPU and desktop NVIDIA GTX 690 GPU card, respectively; the performance by NVIDIA Jetson TK1 is close to that by desktop Intel i7-3770 multi-core CPU, and the desktop NVIDIA GTX 690 GPU has the best performance. From above results, NVIDIA Jetson TK1 is a comparable embedded and mobile board by comparing with other embedded boards and desktop CPUs; moreover, NVIDIA Jetson TK1 has the low cost and low power consumption advantages by comparing with other desktop GPU cards. Hence, it will become a new research direction to study NVIDIA Jetson TK1 in the high performance computing community fields.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Compute Unified Device Architecture (abbreviate to CUDA, (Nickolls et al., 2008)) could access GPUs and has made the supercomputing available to the mass. In the past, many parallel algorithms, CUDA programs and tools have been developed based on these desktop GPU cards in several high-speed computing fields, such as the meteorology, image and video processing, fluid dynamics simulation, seismic analysis, and etc. For example, in the computational biology, there are several tools, such as MUMmerGPU (Schatz et al., 2007), CUDA-MEME (Liu et al., 2010), CUDA-BLASTP (Liu, Schmidt, Mu¨ller-Wittig, 2011), and etc. ClustalW (Thompson, Higgins, Gibson, 1994) is a progressive multiple sequence alignment tool, which is the commonly used approach to align a set of sequences by repeatedly aligning pairs of sequences and previously generated alignments. The idea of progressive multiple sequence alignment is to align pairs of sequences according to the orders in the phylogenetic tree which was built by the similarity scores calculated from each pair of sequences. In the previous work, we have proposed a GPU version of ClustalW v2.0.11, called CUDA ClustalW v1.0 (Hung et al., 2015), by using the intra-task parallelization (Liu, Maskell, Schmidt, 2009) on single- and multi-GPUs. CUDA ClustalW v1.0 based on desktop NVIDIA Tesla C2050 GPU card can achieve more than 33x speedups for overall execution time by comparing to ClustalW v2.0.11 based on desktop Intel XEON X5550 CPU of 2.67GHz.

Computer-Aided Drug Design (abbreviated to CADD) (John, 2007) becomes an emerging research field, and it is helpful to improve the efficiencies of drug design and development. By CADD, the potential inhibitors can be found for the target genes or proteins. After that, in general, these potential inhibitors will be used to compare with compound databases, such as ZINC (Irwin and Shoichet, 2005), in order to find other compounds with the similar structure. Therefore, the compound comparison has been an important and commonly-used task for the computational chemistry. Many algorithms, such as LINGO method (Vidal, Thormann, Pons, 2005) and fingerprint methods (Grant et al., 2006), have been proposed to do the compound comparisons in the past. However, it will be time-consuming when comparing with a large amount of compounds. In the previous work, we have proposed a GPU-based parallel algorithm for multiple compound comparisons, called CUDA-MCC (Lin et al., 2014), on single- and multi-GPUs. CUDA-MCC based on desktop NVIDIA Tesla K20m GPU card can achieve 45 times and 391 times faster than its CPU version based on desktop Intel XEON E5-2650 CPU of 2.0 GHz.

Therefore, in this paper, we construct a mobility and acceleration computing platform for bioinformatics with NVIDIA Jetson TK1. Based on this platform, CUDA ClustalW v1.0 and CUDA-MCC were re-ported, called ClustalWtk and MCCtk, and tested in this paper, respectively. From the experimental results, these tools both can achieve 3 and 4 times speedup ratios on single NVIDIA Jetson TK1 by comparing with their CPU versions on single Intel XEON E5-2650 CPU or single ARM Cortex-A15 CPU, respectively. Moreover, the cost-performance ratio by NVIDIA Jetson TK1 is higher than that by the desktop NVIDIA Tesla K20m GPU card. The user friendly interface can allow users to easily access GPU computational power of the proposed tool; moreover, NVIDIA Jetson TK1 contains Bluetooth and Wi-Fi modules for mobile applications. Hence, the web and mobile services for these two tools with user friendly interfaces also were provided in this work.

The rest of this paper is organized as follows. Preliminary concepts are introduced in Section 2. Section 3 presents the bioinformatics platform constructed on NVIDIA Jetson TK1 and the details of ClustalWtk and MCCtk tools. Experimental results are given in Section 4. Conclusions are drawn in Section 5.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. PRELIMINARY CONCEPTS

2.1. CUDA and NVIDIA GPU Architecture

CUDA is an extension of commonly-used programming languages, such as C/C++, in which users can write multithreading programs for practical applications. The CUDA program is implemented in two parts: host executed by CPU and device executed by GPU. The function executed on the device part is called a kernel, which is invoked as a set of concurrently executing threads. These threads are grouped into a hierarchical architecture: thread blocks and grid. A grid is a set of thread blocks, and a thread block contains many threads. The size of grid is the number of thread blocks per-grid, and the size of thread block is the number of threads per-thread block. Threads in a thread block can communicate through a per-thread block shared memory and synchronize with each other. In addition to shared memory, there are five memory types, includes per-grid private local memory, global memory for data shared by all thread blocks, texture memory, constant memory, and per-thread private registers. The basic processing unit in NVIDIA's GPU architecture is called the streaming processor or CUDA core. Many streaming processors perform the computations on GPU. Several streaming processors can be integrated into a streaming multiprocessor according to various architectures, such as 32 and 192 streaming processors per-streaming multiprocessor for Fermi and Kepler architectures, respectively. The caches of texture memory and constant memory are limited to 8KB per-streaming multiprocessor. Fermi and Kepler architectures have real configurable L1 per-streaming multiprocessor and unified L2 caches among streaming multiprocessors. Hence, L2 caches can be accessed by the global memory and each streaming multiprocessor can use the L1 caches and shared memory. While the program runs the kernel, the device schedules thread blocks for the execution on the streaming multiprocessor. The SIMT scheme refers to threads running on the streaming multiprocessor in a small group of 32, called a warp. The warp scheduler simultaneously schedules and dispatches instructions. The performance by a CUDA program is deeply influenced by the number of threads per-thread block and the number of thread block per-grid.

2.2. ClustalW

ClustalW (Thompson, Higgins, Gibson, 1994) is a progressive multiple sequence alignment tool; and its procedure can be divided into three steps: (1) the distance matrix calculation, (2) the guide tree creation, and (3) the progressive alignment. In most of cases, the computation time of distance matrix calculation step occupied most of overall execution time by ClustalW (Li, 2003). In the distance matrix calculation step, the similarity scores were calculated from each pair of biological sequences by using the pairwise alignment algorithm, such as Needleman and Wunsch algorithm (Needleman and Wunsch, 1970). For aligning two sequences both with length of n, the time and space complexity both is $O(n^2)$ by using Needleman and Wunsch algorithm. For aligning k sequences with length of n by ClustalW, it needs to do $k^2/2$ pairwise alignments and the total time complexity is $O(k^2n^2)$. In the guide tree creation step, a rooted phylogenetic tree was built according to the calculated distance matrix (similarity scores in step1) by using the phylogenetic tree construction algorithm. The time complexity of this step is $O(k^3)$ for most of phylogenetic tree construction algorithms, such as UPGMA (Sneath and Sokal 1973). The built guide tree could be used to decide the orders of following progressive alignment step. In the progressive alignment step, the pairwise alignment results were generated according to the orders in the guide tree (created in step2). Although this step still needs to use the pairwise alignment algorithm to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

obtain the pairwise alignment results, and then combines the previously generated alignments, the times of pairwise alignments in this step generally is far less than that in the distance matrix calculation step.

2.3. Multiple Compound Comparisons (MCC)

In general, molecules (compounds) can be represented as fingerprints (Grant et al., 2006) and SMILES (Vidal, Thormann, Pons, 2005), and then the work of compound comparison can be seen as the string comparison. The Tanimoto coefficient is the most popular measurement between two molecules. For a pairwise compound comparison (denoted as one to one comparison, O2O), the time complexity is $O(n^2)$, where n is the maximal length of compounds. In general, the length of compound is short (tens to hundreds) and the computation time is small. However, it will be time-consuming when comparing with a large amount of compounds. For the MCC problem, it can be done to compare a compound with a set of compounds (denoted as one to all comparisons, abbreviated to O2A) or to compare two sets of compounds (denoted as all to all comparisons, abbreviated to A2A). For A2A, the intrinsic time complexity of MCC problem is $O(k^2n^2)$ with k compounds of length n. The LINGO method (Vidal, Thormann, Pons, 2005) is to model a molecule as a collection of substrings of SMILES representation. Thus, a SMILES string is fragmented into all substrings with length of q by using the sliding window scheme. These substrings are stored as a set of q-Lingos. When doing O2O, two sets of q-Lingos from two molecules are used to find the same Lingos. The number of same Lingos is then used to calculate the Tanimoto coefficients.

3. METHOD

3.1. Constructing Bioinformatics Platform

In this paper, a bioinformatics platform is constructed on NVIDIA Jetson TK1, which can be separated into four parts: *system setting up*, *program porting*, *performance tuning*, and *interface design*. For *system setting up*, some preparation works need to be done, such as refresh system image, set up network and install some software in the proposed computing platform. For *program porting*, for ClustalWtk and MCCtk, the compile commands and program codes of CUDA ClustalW v1.0 and CUDA-MCC should be regulated and modified, respectively. For *performance tuning*, it is necessary due to the different of numbers of thread blocks and threads per-thread block between NVIDIA Jetson TK1 and desktop GPU cards. ClustalWtk and MCCtk will be evaluated on NVIDIA Jetson TK1 by various numbers of thread blocks and the number of threads per thread block, respectively. For *interface design*, various interfaces by Qt, web and mobile services are designed for ClustalWtk and MCCtk, respectively.

3.1.1. System Setting Up

All the works in this paper are implemented on Linux for Tegra (abbreviate to L4T) R21.1. System image and device driver are downloaded from NVIDIA Developer Zone (https://developer.nvidia.com/linux-tegra-rel-21), which provides some useful development tools for NVIDIA. The decompression path of these downloaded files is very important, e.g., for device driver, it must be placed in the child directory of system image, named “rootfs”. It then is further compiled by executing “apply_binaries.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

sh" on the top directory of system image. A micro USB cable is used to be connected between NVIDIA Jetson TK1 and Linux PC-host. To place system in the "Force USB Recovery Mode", the device must be powered off, not suspend or sleep state. Then reset button is pressed while force recovery button is held. To examine whether the device is entered the "Force Recovery Mode", the command "lsusb" is utilized to check if "NVidia Corp" exists or not. If "NVidia Corp" exists, NVIDIA Jetson TK1 has successfully connected with Linux PC-host. In order to flash device, the command "sudo ./flash.sh -S 14580MiB jetson-tk1 mmcblk0p1" is executed. After finished this flashing action, default account and password are used to login the L4T.

NVIDIA Jetson TK1 can be accessed in two possible ways: *direct access* and *remote access*. The first way is using keyboard, mouse and HDMI monitor to directly access to NVIDIA Jetson TK1 in the "GUI mode". The other is using Ethernet port to access to the board (NVIDIA Jetson TK1) in the "remote mode". Since DHCP and SSH both are default "on", NVIDIA Jetson TK1 can be accessed by using SSH in the "remote control mode". IP is assigned by DHCP server when NVIDIA Jetson TK1 is connected to a router. It is important to tell "apt" not to overwrite the file "libglx.so" in L4T R19 if you upgrade the system. It seems to be fixed in L4T R21. "libglx.so" is a specific file in the NVIDIA's graphics driver that might get replaced by an incorrect version from Ubuntu that stops users from being able to boot into the graphical environment. Therefore, it should execute "sudo apt-mark hold xserver-xorg-core" on NVIDIA Jetson TK1 before connecting it to Internet or perform an update.

The version CUDA 6.5 is installed on NVIDIA Jetson TK1. The ".deb" file of the CUDA Toolkit for L4T is also downloaded from NVIDIA Developer Zone. In order to install the CUDA repo metadata, "sudo dpkg -i cuda-repo-l4t-r21.1-6-5-prod_6.5-14_armhf.deb" is executed (downloaded manually for L4T). "sudo apt-get update" is executed to download and install the actual CUDA Toolkit including the OpenGL toolkit from NVIDIA. "sudo apt-get install cuda-toolkit-6-5" is executed to install the package and the account should be added to the "video" group to allow access GPU. It is also necessary to add "export PATH=/usr/local/cuda/bin:$PATH" and "export LD_LIBRARY_PATH=/usr/local/cuda/lib:$LD_LIBRARY_PATH" to "bashrc" file and then to source it.

Other development tools also should be installed. Gcc and G++ are necessary for building OpenMPI and Qt; Apache and PHP are essential for application interfaces design; Gcc, G++, Apache, PHP, mail server, and Qt4 are available on apt package manager which can be directly installed by executing "sudo apt-get install gcc g++ apache2 php5 libapache2-mod-php5 mailutils qt4-dev-tools libqt4-dev libqt4-core libqt4-gui libqt4-opengl" on terminal (Qt: https://devtalk.nvidia.com/default/topic/752892/qt-anyone-/; mailutils: http://askubuntu.com/questions/12917/how-to-send-mail-from-the-command-line; Apache: https://www.digitalocean.com/community/tutorials/how-to-install-linux-apache-mysql-php-lamp-stack-on-ubuntu). OpenMPI is download from the OpenMPI website (http://www.open-mpi.org/) as ".tar.gz" file. The OpenMPI source file is decompressed and configured with the building path: "—prefix=/usr/local/openmpi". After "make all" and "sudo make install" have finished, "export OPENMPI=/usr/local/openmpi", "export LD_LIBRARY_PATH=$OPENMPI/lib" and "export PATH=$OPENMPI/bin:$PATH" should add to the "bashrc" file and then to source it.

A Wi-Fi module (Intel 7260 mini PCIE) is installed on the NVIDIA Jetson TK1; however, the system kernel seems not support this module. Several reports (https://github.com/jetsonhacks/install7260LT4211) have been tried to fix this problem. Device driver for the Intel 7260 miniPCIE card is compiled for the NVIDIA Jetson TK1 under LT4 R21.1, which is available from GitHub website (https://github.com/). To avoid crashes, "/etc/rc.local" and the "cfg80211.ko" and "mac80211.ko" files are backed before proceeding. After "install7260Driver.sh" has been executed, the Wi-Fi module installation is successful.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.1.2. Program Porting

In this part, the program codes of ClustalWtk and MCCtk are modified from CUDA ClustalW v1.0 and CUDA-MCC designed for the desktop GPU cards, and then they both were compiled on NVIDIA Jetson TK1, respectively. The details of ClustalWtk and MCCtk are described in the Section 3.2 and Section 3.3, respectively. Some errors have been reported in the compiler, such as: "library could not found. Fixing Configure and Makefile is necessary". In this case, the library path for CUDA should be forced to use x86 path, and device compute capability should be set at 3.2. For program with OpenMP, "-Xcompiler -fopenmp" is added to the compile command. The parameter "-O3" is also attached to achieve the better performance.

However, the running time is incorrect in NVIDIA Jetson TK1 when the input data is large. Many experiments have showed that the device would return tiny or negative running time when it is longer than two thousand seconds. This situation is not appeared on the desktop CPU. The running time is computed by clock() function and stored in the type clock_t. According to the definition in "time.h" file, type "clock_t" is type "long". The maximum of type "long" on NVIDIA Jetson TK1 is 2147483647, and on the desktop CPU, such as Intel XEON E5-2650, is 9223372036854775807. The maximum of type "long" on NVIDIA Jetson TK1 is divided by CLOCKS_PER_SEC (1000000) is about 2147, in line with the aforementioned situation. Since the clock() function will cause the overflow, the "ftime()" function and "timeb" structure are the better choice. The "timeb" structure consists of "time_t" time and unsigned short "millitm". The running time divided into two parts for the storage can be calculated for the longer time.

3.1.3. Performance Tuning

For ClustalWtk and MCCtk, *performance tuning* is further divided into two parts: *number of thread blocks/threads per-thread block* and *hardware control*. For the first part, when the numbers of thread blocks and threads per-thread block is set at 128 and 64, respectively, the best performance by ClustalWtk or MCCtk is obtained. These numbers set for ClustalWtk and MCCtk are different to those of CUDA ClustalW v1.0 and CUDA-MCC. For the second part, "cat /sys/devices/system/cpu/online" is used to check the status of CPUs. Since only one CPU core is default on in NVIDIA Jetson TK1, other CPUs should be turned on by hand. For example, CPU2 is turned on by "echo 1 > /sys/devices/system/cpu/cpu2/online". According to the consumption of battery (power), the modes of CPU are separated into three types: "quiet", "normal" and "performance". To obtain the best power efficiency, default "quiet" mode is turned off to enter the "performance" mode. "echo 0 > /sys/devices/system/cpu/cpuquiet/tegra_cpuquiet/enable" is used to turn off the "quiet" mode. "echo performance > /sys/devices/system/cpu/cpu2/cpufreq/scaling_governor" is used to keep frequency on static instead of a range. CPU frequency can be set by writing this: "echo <*frequency*> > /sys/devices/system/cpu/cpu0/cpufreq/scaling_setspeed". If the type is running on the "quiet" mode, the following frequencies (Hz) are available: 51000, 102000, 204000, 312000, 564000, 696000, 828000, 960000, 1092000. And if the type is running on the "performance" mode, the following frequencies (Hz) are available: 204000, 312000, 564000, 696000, 828000, 960000, 1092000, 1122000, 1224000, 1326000, 1428000, 1530000, 1632000, 1734000, 1836000, 1938000, 2014500, 2116500, 2218500, 2320500. To manually control the clock frequencies (Hz) of the GPU, supported rates should be first determined by "cat /sys/kernel/debug/clock/gbus/possible_rates". Then

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

set a rate (e.g. 852000 kHz) by "echo 852000000 > /sys/kernel/debug/clock/override.gbus/rate" and "echo 1 > /sys/kernel/debug/clock/override.gbus/state". Finally, the rate is verified by "cat /sys/kernel/debug/clock/gbus/rate".

3.1.4. Interface Design

The user friendly interfaces are designed for ClustalWtk and MCCtk by Qt and Web, respectively. Qt is accessed through SSH and X11 forwarding and only available on PC-host. Web is accessed through IP address and has more device compatibility.

The stable version Qt 4.8 has more materials on internet or bookstore. Qt application is accessible to operating CUDA program through SSH and X11 forwarding. X11 forwarding is default on when NVIDIA Jetson TK1 is out of the factory. "ssh -X <*IP*>" can be used to connect device with X11 forwarding on Linux or Mac, and "Xming" and "Putty" application are the solution for Microsoft Windows. "QFile" is used to read and write file on NVIDIA Jetson TK1, and "QProcess" is used to start external programs and to communicate with them. The CUDA program is executed by "start()" function. "Connect()" function is used to report the state while CUDA program has finished. In "start()" function, bash command is invalid since the system environment is not work in it. Absolute path of program is added instead of using "$HOME" variable or relative path. In the end, bash command "mail -s <*subject*> -A <*attached file*>" is added in the "system()" function to send result data to users. No permission problem will be appeared if the account (ubuntu) has been added to "video" group.

Bootstrap is employed as the website framework for popular feature in developing responsive, mobile first projects on the web. It can be applied on the mobile phone, pad and PC. HTML provides users to set up input file and parameters, and send it to PHP script by AJAX techniques. According to user's choice, PHP script executes CUDA program in CPU or GPU version. The running time of CUDA program, such as ClustalWtk or MCCtk, is averaged more than 30 seconds, largely depending on input data and algorithm. As the default running time for PHP script is limited to 30 seconds, it is necessary to set the max execution time value before running. "set_time_limit(0)" is added to PHP script, and zero means unlimited. The "shell_exec()" function is utilized to execute CUDA program on PHP script. "cuda device not found" are reported when PHP runs CUDA program. The main reason is that apache account (www-data in ubuntu/L4T) does not have permission to access GPU. Hence, "www-data" is added to "video" group and it work perfectly.

3.2. ClustalWtk

In ClustalWtk, the distance matrix calculation step is implemented on GPU and several optimization methods were designed to enhance the performance by ClustalWtk. In the distance matrix calculation step, the similarity scores were calculated from each pair of sequences by using pairwise alignment algorithm. For *k* sequences, ClustalWtk used the intra-task parallelization to assign a pairwise alignment to a thread block, which each pairwise alignment is implemented by the *Synchronous Diagonal Multiple Threads* (*SDMT*) type (Lin and Lin, 2014). The pairwise alignment algorithm used in ClustalWtk includes three paths: forward path, reverse path, divide and conquer path. In order to do the computations in the forward array and the reverse array concurrently, the threads in a thread block are divided into two parts, one for forward array and another for reverse array. Only one thread in a thread block is then used to merge the scores in the last row of forward array and reverse array in order to find the break point

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and then divide the original two-dimensional array into two sub-arrays. The divide and conquer path is done by the loop function. Besides, in ClustalWtk, several optimization methods below were designed for the distance matrix calculation step:

- **Load balancing strategy:** Using the intra-task parallelization, a pairwise alignment is assigned to a thread block. Although there is only one streaming multiprocessor in NVIDIA Jetson TK1, a streaming multiprocessor can deal with one or more thread block(s). For k sequences in ClustalWtk, a kernel function only deals with a fixed number of pairwise alignments with the similar length of sequences. The number of threads per-thread block and the number of thread block per-grid in ClustalWtk are 64 and 128, respectively, according to Section 3.1.3;
- **Memory allocation:** For a pairwise alignment, the inputs are two sequences and a substitution matrix. These two sequences are stored in the shared memory and the substitution matrix is stored in the texture memory. All of sequences and the distance matrix are stored in the 2GB RAM;
- **Unsigned char data type:** Since each pair of sequences stored in the shared memory, the shared memory usage will affect the occupancy of streaming multiprocessors. These two sequences are stored by using the *unsigned char* data type, not *int*, in order to reduce the shared memory usage and increase the occupancy;
- **Precision adjustment:** Since the single-precision floating-point operations in CUDA are not designed according to the IEEE-754 standard. In ClustalWtk, *__fdiv_rn*() function was used to replace the built-in division operator.

3.3. MCCtk

In MCCtk, the goal is to compare two sets of compounds (A2A) listed as Query and Database at first, and then find the compounds in Database with more than 0.85 Tanimoto coefficient for each compound in Query. There are three phases in MCCtk: *preprocessing*, *comparison* and *output*. In the *preprocessing* phase, the compounds in Query and Database are stored as a two dimensional string arrays (represented as the SMILES code), *Q* and *Db*, respectively. After that, each SMILES code (a compound) is then fragmented into a set of 4-Lingos (as sub-strings) by using the sliding window scheme with an offset 1. For each 4-Lingo, it will be transformed into a 32-bit integer (called LINGO score) according to the ASCII code table. The LINGO length, LINGO number and LINGO magnitude are calculated, respectively, according to the number of 4-Lingos, the times of each 4-Lingo appeared in the SMILES code and the number of 4-Lingos without repeats. For the *comparison* phase, a compound in Query will be used to compare with all compounds in Database by all threads (the inter-task parallization) when executing the kernel function once. In a thread block, each thread is used to compare a compound in Query with one of compounds in Database. Hence, the computing workload of each thread in a thread block should be equal in order to achieve high performance. A load-balancing strategy can be used to accelerate the computation speed on NVIDIA Jetson TK1 by considering the LINGO score, LINGO length, LINGO number, and LINGO magnitude. After the *comparison phase*, in the *output* phase, for each compound in Query, the compounds in Database with more than 0.85 Tanimoto coefficients are reported.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. EXPERIMENTAL RESULTS

In this work, ClustalWtk and MCCtk were implemented by C+CUDA on single NVIDIA Jetson TK1. In order to evaluate these two tools, a machine and a NVIDIA Jetson TK1 are used in the experimental tests. This machine has eight Intel XEON CPU cores, four GPU cards and 128GB RAM. Each Intel XEON CPU core is 4 Intel XEON E5-2650 processors of 2.0 GHz and each GPU card is NVIDIA Tesla K20m with 4.8GB RAM and 13 streaming multiprocessors, each multiprocessor has 192 CUDA cores of 0.706 GHz. The NVIDIA Jetson TK1 has 1 ARM Cortex CPU core, 1 Tegra K1 Soc and 2GB RAM. The ARM Cortex CPU core is 4 Cortex-A15 processors of 2.32GHz and the Tegra K1 is NVIDIA Kepler GK20a with 1 streaming multiprocessor, which has 192 CUDA cores of 0.852 GHz. ClustalW v2.0.11 and MCC are run based on Intel XEON E5-2650 CPU in the machine and ARM Cortex-A15 CPU in NVIDIA Jetson TK1, respectively; CUDA ClustalW v1.0 and CUDA-MCC are run based on NVIDIA Tesla K20m GPU card in the machine; ClustalWtk and MCCtk are run based on NVIDIA Kepler GK20a in NVIDIA Jetson TK1.

For the evaluation of ClustalW v2.0.11, CUDA ClustalW v1.0 and ClustalWtk, the test protein sequences (C.-L. Hung et al., 2015) were downloaded from the NCBI website (www.ncbi.nlm.nih.gov/), and these sequences can be classified into eight test sets: (c1) 100 sequences with length of 97, (c2) 100 sequences with length of 498, (c3) 100 sequences with length of 1002, (c4) 100 sequences with length of 1523, (c5) 1000 sequences with length of 97, (c6) 1000 sequences with length of 498, (c7) 1000 sequences with length of 1002, (c8) 1000 sequences with length of 1523. For the evaluation of MCC, CUDA-MCC and MCCtk, the test compounds are randomly selected from the ZINC database [14] and then these compounds are divided into Query and Database sets. There are three test sets used in the following experiment tests: (m1) 4096 compounds in Query and Database, respectively, (m2) 8192 compounds in Query and Database, respectively, (m3) 16384 compounds in Query and Database, respectively.

The first test is to evaluate ClustalW v2.0.11, CUDA ClustalW v1.0 and ClustalWtk for these eight test sets on the machine and single NVIDIA Jetson TK1, respectively. Figure 1(a) shows the execution time (unit: second), which is represented by $\log_{10}$, by ClustalW v2.0.11, CUDA ClustalW v1.0 and ClustalWtk on various test sets, respectively. From Figure 1(a), we can see that the execution time by CUDA ClustalW v1.0 on single NVIDIA Tesla K20m GPU card and that by ClustalWtk on single NVIDIA Jetson TK1 all is less than that by ClustalW v2.0.11 on single Intel XEON E5-2650 CPU and ARM Cortex-A15 CPU, respectively. ClustalWtk can achieve 3 times speedup ratio by comparing the execution time on single NVIDIA Jetson TK1 with that by ClustalW v2.0.11 on single Intel XEON E5-2650 CPU. ClustalWtk can achieve 4 times speedup ratio by comparing the execution time on single NVIDIA Jetson TK1 with that by ClustalW v2.0.11 on single ARM Cortex-A15 CPU. The execution time by CUDA ClustalW v1.0 on single NVIDIA Tesla K20m GPU card is less than that by ClustalWtk on single NVIDIA Jetson TK1. The reason is that the number of CUDA cores on single NVIDIA Tesla K20m GPU card is 13 times than that on single NVIDIA Jetson TK1. However, the execution time by ClustalWtk on single NVIDIA Jetson TK1 is only 5 times than that by CUDA ClustalW v1.0 on single NVIDIA Tesla K20m GPU card. Hence, in ClustalWtk, the cost-performance ratio by NVIDIA Jetson TK1 is better than that by NVIDIA Tesla K20m GPU card.

The second test is to evaluate MCC, CUDA-MCC and MCCtk for these three test sets on the machine and single NVIDIA Jetson TK1, respectively. Figure 1(b) shows the execution time (unit: second), which is represented by $\log_{10}$, by MCC, CUDA-MCC and MCCtk on various test sets. From Figure 1(b), we can see that the execution time by CUDA-MCC on single NVIDIA Tesla K20m GPU card and that by

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MCCtk on single NVIDIA Jetson TK1 all is less than that by MCC on single Intel XEON E5-2650 CPU and ARM Cortex-A15 CPU, respectively. MCCtk can achieve 3 times speedup ratio by comparing the execution time on single NVIDIA Jetson TK1 with that by MCC on single Intel XEON E5-2650 CPU. MCCtk can achieve 4 times speedup ratio by comparing the execution time on single NVIDIA Jetson TK1 with that by MCC on single ARM Cortex-A15 CPU. In Figure 1(b), for the comparisons of NVIDIA Tesla K20m GPU card and NVIDIA Jetson TK1, we have the same observations from the Figure 1(a). In MCCtk, the cost-performance ratio by NVIDIA Jetson TK1 is also better than that by NVIDIA Tesla K20m GPU card.

User friendly interfaces designed by Qt for ClustalWtk and MCCtk are shown in Figures 2(a) and 2(b), respectively. For ClustalWtk, Figure 2(a-i) is the initial interface for ClustalWtk and the input "biological sequences" can be pasted on the interface; the input also can be selected from the directory as shown in Figure 2(a-ii); one of two tools, ClustalW runs on desktop CPUs and ClustalWtk runs on NVIDIA Jetson TK1, can be selected by users to obtain the alignment results as shown in Figure 2(a-i); the execution time by ClustalW or ClustalWtk can be listed in the interface as shown in Figure 2(a-i); the output "multiple sequence alignment results" can be shown in the interface as shown in Figure 2(a-iii). For MCCtk in the left side of Figure 2(b), the Query and Database files can be selected from the directory; the compound in Query also can be assigned in the interface; the other compounds with similar structure are calculated by MCCtk and the information of these compounds is shown. The 3D structure visualization of selected compound can be shown in the right side of Figure 2(b).

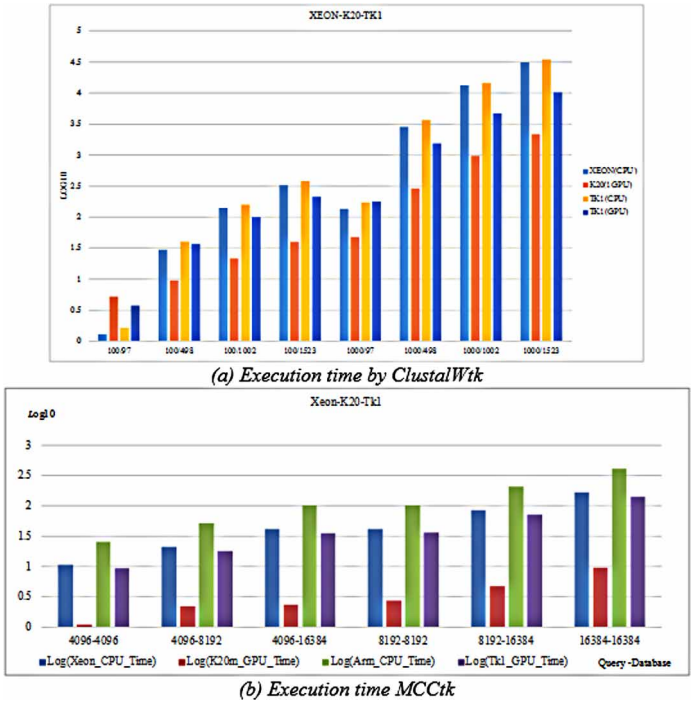
The web and mobile services also can be designed in the NVIDIA Jetson TK1 by Ethernet and Wi-Fi modules. The interfaces of the web and mobile services by ClustalWtk and MCCtk are shown in Figures 3 and 4, respectively. For ClustalWtk, Figure 3(a-i) is the initial interface on the web service; the input sequences can be selected or pasted on the web as shown in Figure 3(a-ii); the one of two tools, ClustalW and ClustalWtk, can be selected to obtain the results of multiple sequence alignment as shown in Figure 3(a-ii); the execution time by ClustalW or ClustalWtk can be displayed in the Figure 3(a-iii); the results are shown in Figure 3(a-iv). Figure 3(b-i) is the initial interface on the mobile service; the input sequences can be selected and then be uploaded to NVIDIA Jetson TK1 as shown in Figure 3(b-ii); the output is displayed in the Figure 3(b-iii). For MCCtk, Figure 4(a-i) is the initial interface on the web service; the inputs, Query and Database files, can be selected on the web as shown in Figure 4(a-ii); the information of compounds with similar structure are listed in Figure 4(a-iii). Figure 4(b-i) is the initial interface on the mobile service; the execution time by MCCtk is shown in Figure 4(b-ii); the information of resulted compounds are listed in Figure 4(b-iii).

5. CONCLUSION

In this paper, a mobility and acceleration computing platform for bioinformatics by NVIDIA Jetson TK1 was constructed. Moreover, ClustalWtk and MCCtk were designed based on the platform. From the experimental results, two observations for ClustalWtk and MCCtk were summarized in this work: (1) they both achieved 3 and 4 times speedup ratios on single NVIDIA Jetson TK1 by comparing with ClustalW v2.0.11 and MCC on single Intel XEON E5-2650 CPU and ARM Cortex-A15 CPU, respectively. (2) The cost-performance ratio by NVIDIA Jetson TK1 is better than that by NVIDIA Tesla K20m GPU card. The user friendly interfaces of Qt, web and mobile services were designed for ClustalWtk and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

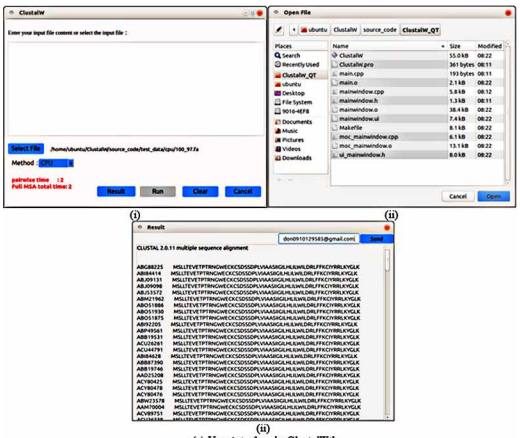
Figure 1. The execution time by ClustalWtk and MCCtk (unit: second, represented by log_{10})



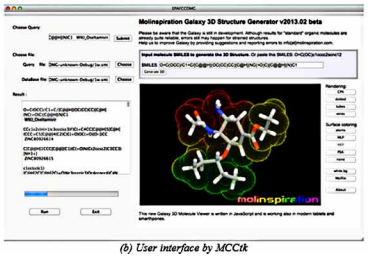
(a) Execution time by ClustalWtk

(b) Execution time MCCtk

Figure 2. User friendly interfaces with input and output procedures proposed by ClustalWtk and MCCtk



(a) User interface by ClustalWtk



(b) User interface by MCCtk

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Web and mobile services by ClustalWtk

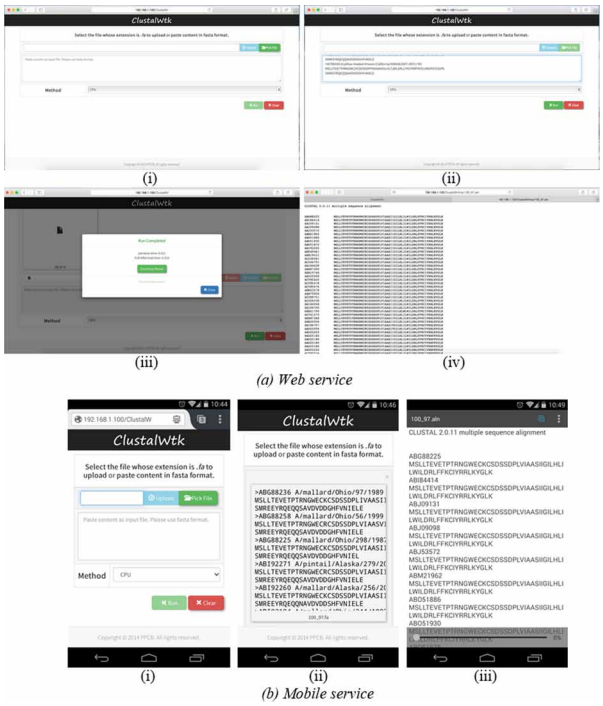
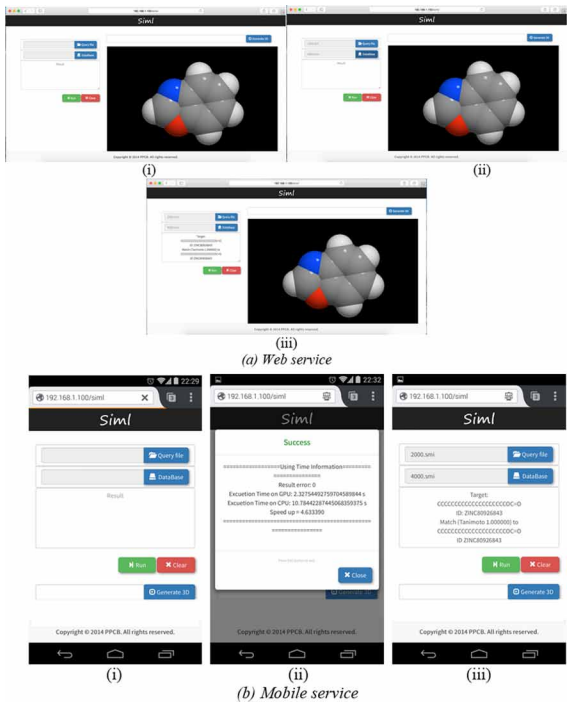


Figure 4. Web and mobile services by MCCtk



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MCCtk, respectively. In the future, several parallel techniques, such as thread-level parallelism (Wu et al., 2014) and multi-GPU communication (Playne, Hawick, 2015), will be considered to apply into the proposed platform and tools.

ACKNOWLEDGMENT

Part of this work was supported by the Ministry of Science and Technology under the grants MOST 103-2632-E-126-001-MY3, MOST 104-2221-E-182-050 and MOST 104-2221-E-182-051.

REFERENCES

Apache. (n. d.). Retrieved from https://www.digitalocean.com/community/tutorials/how-to-install-linux-apache-mysql-php-lamp-stack-on-ubuntu

Chang, S. Fu, R., Couture, S., Menarini, M., Escobar, M. A., Kuteifan, M., Lubarda, M., Gabay, D., & Lomakin, V. (2015). Micromagnetics on high-performance workstation and mobile computational platforms. *Journal of Applied Physics*, 117(17). doi:10.1063/1.4918638

Grant, J. A., Haigh, J. A., Pickup, B. T., Nicholls, A., & Sayle, R. A. (2006). Lingos, Finite State Machines, and Fast Similarity Searching. *Journal of Chemical Information and Modeling*, *46*(5), 1912–1918. doi:10.1021/ci6002152 PMID:16995721

Hung, C.-L., Lin, Y.-S., Lin, C.-Y., Chung, Y.-C., & Chung, Y.-F. (2015). CUDA ClustalW: An efficient parallel algorithm for progressive multiple sequence alignment on Multi-GPUs. *Computational Biology and Chemistry*, *58*, 62–68. doi:10.1016/j.compbiolchem.2015.05.004 PMID:26052076

Irwin, J. J., & Shoichet, B. K. (2005). ZINC-A Free Database of Commercially Available Compounds for Virtual Screening. *Journal of Chemical Information and Modeling*, *45*(1), 177–182. doi:10.1021/ci049714+ PMID:15667143

Jetson TK1. (n. d.). Retrieved from http://elinux.org/Jetson_TK1

John, H. V. D. (2007). Computer-aided drug design: The next 20 years. *Journal of Computer-Aided Molecular Design*, *21*, 1591–1601. PMID:17989929

Li, K. B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics (Oxford, England)*, *19*(12), 1585–1586. doi:10.1093/bioinformatics/btg192 PMID:12912844

Lin, C. Y., & Lin, Y. S. (2014). Efficient parallel algorithm for multiple sequence alignments with regular expression constraints on graphics processing units. *International Journal on Computer Science and Engineering*, *9*(1/2), 11–20. doi:10.1504/IJCSE.2014.058687

Lin, C. Y., Wang, C.-H., Hung, C.-L., & Lin, Y.-S. (2014). Efficient Parallel Algorithm for Compound Comparisons on Multi-GPUs. *Proc. The first International Workshop on High Performance Computing on Bioinformatics* (pp. 30-35). 10.1109/BIBM.2014.6999307

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, W., Schmidt, B., & Müller-Wittig, W. (2011). CUDA-BLASTP: Accelerating BLASTP on CUDA-Enabled Graphics Hardware. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(6), 1678–1684. doi:10.1109/TCBB.2011.33 PMID:21339531

Liu, Y., Maskell, D. L., & Schmidt, B. (2009). CUDASW++: Optimizing Smith-Waterman sequence database searches for CUDA-enabled graphics processing units. *BMC Research Notes*, *2*(1), 73. doi:10.1186/1756-0500-2-73 PMID:19416548

Liu, Y., Schmidt, B., Liu, W., & Maskell, D. L. (2010). CUDA–MEME: Accelerating motif discovery in biological sequences using CUDA-enabled graphics processing units. *Pattern Recognition Letters*, *31*(14), 2170–2177. doi:10.1016/j.patrec.2009.10.009

Mailutils. (n. d.). Retrieved from http://askubuntu.com/questions/12917/how-to-send-mail-from-the-command-line

NCBI. (n. d.). Retrieved from www.ncbi.nlm.nih.gov/

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, *48*(3), pp443–pp453. doi:10.1016/0022-2836(70)90057-4 PMID:5420325

Nickolls, J., Buck, I., Garland, M., & Skadron, K. (2008). Scalable parallel programming with CUDA. *ACM Queue; Tomorrow's Computing Today*, *6*(2), 40–53. doi:10.1145/1365490.1365500

NVIDIA Developer Zone. (n. d.). Retrieved from https://developer.nvidia.com/linux-tegra-rel-21

NVIDIA Jetson TK1 Development Kit. (2014, April). Retrieved from http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html

OpenMPI. (n. d.). Retrieved from http://www.open-mpi.org/

Paolucci, P. S., Ammendola, R., Biagioni, A., Frezza, O., Cicero, F. L., Lonardo, A., . . . Vicini, P. (2015). *Power, Energy and Speed of Embedded and Server Multi-Cores applied to Distributed Simulation of Spiking Neural Networks: ARM in NVIDIA Tegra vs Intel Xeon quad-cores*. arXiv:1505.03015.

Playne, D. P., & Hawick, K. A. (2015). Benchmarking multi-GPU communication using the shallow water equations. *Int. J. of Big Data Intelligence*, *2*(3), 157–167. doi:10.1504/IJBDI.2015.070596

Qt. (n. d.). Retrieved from https://devtalk.nvidia.com/default/topic/752892/qt-anyone-/

Schatz, M. C., Trapnell, C., Delcher, A. L., & Varshney, A. (2007). High-throughput sequence alignment using Graphics Processing Units. *BMC Bioinformatics*, *8*(1), 474. doi:10.1186/1471-2105-8-474 PMID:18070356

Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical Taxonomy*. San Francisco, CA.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. doi:10.1093/nar/22.22.4673 PMID:7984417

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Vidal, D., Thormann, M., & Pons, M. P. (2005). LINGO, an Efficient Holographic Text Based Method To Calculate Biophysical Properties and Intermolecular Similarities. *Journal of Chemical Information and Modeling*, *45*(2), 386–393. doi:10.1021/ci0496797 PMID:15807504

Wolfer, J. (2015). A Heterogeneous Supercomputer Model for High-Performance Parallel Computing Pedagogy. *Proc. IEEE Global Engineering Education Conference* (pp. 799-805). 10.1109/EDUCON.2015.7096063

Wu, C.-C., Ke, J.-Y., Lin, H., & Jhan, S.-S. (2014). Adjusting Thread Parallelism Dynamically to Accelerate Dynamic Programming with Irregular Workload Distribution on GPGPUs. *International Journal of Grid and High Performance Computing*, *6*(1), 1–20. doi:10.4018/ijghpc.2014010101

This research was previously published in the International Journal of Grid and High Performance Computing (IJGHPC), 7(4); edited by Emmanuel Udoh, Ching-Hsien Hsu, and Mohammad Khan; pages 57-73; copyright year 2015 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 36
Genetic Diagnosis of Cancer by Evolutionary Fuzzy-Rough based Neural-Network Ensemble

Sujata Dash
North Orissa University, India

Bichitrananda Patra
KMBB College of Engineering and Technology, India

ABSTRACT

High dimension and small sample size is an inherent problem of gene expression datasets which makes the analysis process more complex. The present study has developed a novel learning scheme that encapsulates a hybrid evolutionary fuzzy-rough feature selection model with an adaptive neural net ensemble. Fuzzy-rough method deals with uncertainty and impreciseness of real valued gene expression dataset and evolutionary search concept optimizes the subset selection process. The efficiency of the hybrid-FRGSNN model is evaluated by the proposed neural net ensemble learning algorithm. Again to prove the learning capability of ensemble algorithm, performance of the component classifiers pairing with FR, GSNN and FRGSNN are compared with proposed hybrid-FRGSNN based ensemble model. In addition to this, efficiency of neural net ensemble is compared with two classical and one advanced ensemble learning algorithms.

1. INTRODUCTION

Microarray technology produces high dimensional datasets by measuring the expression levels of tens of thousands of genes in a single experiment under varying conditions. It has become an indispensable tool for biological, medical and pharmaceutical researchers to get a better understanding of the diseases at genomic level. On the other hand, the inherent problem i.e., large number of features and small sample

DOI: 10.4018/978-1-7998-1204-3.ch036

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

size of microarray dataset makes the analysis process difficult for the problem. Typically, a relatively small number of features are found to be strongly correlated with the phenotypes in question. Therefore, to identify these discriminative features from gene expression dataset, a data mining tool (Witten, Frank & Hall, 2011) known as feature selection technique plays an important role.

The methods specifically used for feature selection can be categorized into two major groups namely, filter and wrapper methods (Saeys, Inza & Larranaga, 2007; Guyon, Nikravesh, Zadeh, 2006). Filter methods select features considering individual characteristic of each feature without taking into account the mutual dependencies among features. Then the features are sorted by their assigned ranks. The top ranked features are kept for further analysis by removing low ranked features. Actually, these selected features are used to develop the diagnostic model to efficiently predict the diseases. On the contrary, in wrapper methods a search algorithm is wrapped around a learning algorithm: so that an estimated learning accuracy for all subsets can be calculated to derive an optimal one. This method is computationally intensive in comparison to filter methods because to obtain an optimal subset of features all possible subsets need to be examined which is practically a difficult task. Thus, to alleviate this difficulty in wrapper methods an intelligent search strategy (Ghosh & Jain (Eds), 2005; Akadi, Amine, Ouardighi & Aboutajdine, 2009; Dash, 2016) may be adopted from a set of heuristic or stochastic algorithms which can be able to generate an optimal subset effectively.

Rough set theory (RST) was introduced by Pawlak (1982) and has been used widely by researchers as a classifier and selection technique. Primarily, it removes redundant data (Dash & Patra, 2012) while keeping intact the characteristics of the problem for classification. The main drawback of rough set (Dash & Patra, 2013) is it cannot handle real valued data whereas microarray dataset contains all real valued gene expression data. In this case, the expression values of genes are discredited before being used and the dataset loses some information during this process which makes the dataset inadequate for analysis purpose. The best alternative to get rid of this problem is either using real valued rough set or fuzzy-rough set theory (Zahi, 2011; Jensen & Shen, 2004; Dash, 2015). Both of them handle the concept of vagueness and indiscernibility which is prevalent in most real life problems (Xu, Miao & Wei, 2009) by complementing each other.

The ensemble learning approach constructs several classifier models for the original dataset and then combines the predictive outputs to identify an unknown sample. The motivation of combining several classifiers is to improve the classification efficiency which in turn depends on the accuracy and diversity (Yang P., Yang H., Bing, Zomaya, 2010) of the base classifiers. The ensemble technique is very popular in the field of classification and pattern recognition as it increases the generalization and percentage of classification by aggregating (Chen, Hong, Deng, Yang, Wei & Cui, 2015) the outcome of finite number of neural network classifiers (Lee, Hong & Kim, 2009a). However, neural network ensemble learning has been used in many problems, such as, face recognition (Lee, Hong & Kim, 2009 b), digital image processing (Liu, Cui, Jiang & Ma, 2004) and medical diagnosis (Huang, Zhou, Zhang & Chen, 2000) and has given outstanding performance in terms of classification accuracy.

In this study, a hybrid gene selection model (Dash & Patra, 2012) is introduced by hybridizing fuzzy-rough filter model (FR) with an evolutionary genetic search based wrapper model (GSNN) (Ghosh & Jain (Eds), 2005). The efficiency of the hybrid model (FRGSNN) (Yang, Chuang & Yang, 2009) is evaluated by the proposed adaptive neural net ensemble learning algorithm. Again to prove the learning capability of ensemble algorithm, performance of the component classifiers pairing with FR, GSNN and FRGSNN are compared with proposed hybrid FRGSNN based ensemble model. In addition to this, efficiency of neural net ensemble is compared with two ensemble learning algorithms.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The paper is structured in the following way: section 2 gives an outline of the proposed learning scheme followed by section 3 furnishing the background of rough set theory, fuzzy-rough set and genetic algorithm. Section 4 and 5 illustrate the proposed learning model and all eight microarray datasets respectively. Experimental setup, results, analysis of comparison of the model, discussion and future scope are presented in section 6. Section 7 presents concluding remarks followed by exhaustive references.

2. AN OUTLINE OF THE LEARNING SCHEME

In this section, an outline of the learning scheme is presented. For a given gene expression dataset, DATA, the proposed learning scheme will work as follows:

Step 1: Replace all missing values with the means of the dataset.
Step 2: *[Filter Method, FR]* Fragment dataset into lower and upper approximations using rough set theory; *DATA → (DATA+) U (DATA -).*
Step 3: Apply fuzzy-rough feature selection method to remove redundant features.
Step 4: Measure the quality of the features using weak gamma evaluator in fuzzy rough lower approximation.
Step 5: [*Wrapper Search, GSNN*] Use reduct sets as the candidate solution of the initial population of genetic search method.
Step 6: Construct a fitness function using KNN learning algorithm and evaluate the fitness values of current population.
Step 7: Chose two fittest chromosomes from the population and apply crossover and mutation to produce new population.
Step 8: Estimate the fitness function for new population and continue the process as long as the value of fitness function increases or the termination condition is satisfied.
Step 9: *[Hybrid search FRGSNN]* Select the optimal subset of features combining filter (FR) and wrapper (GSNN) search.
Step 10: Estimate effectiveness of the selected discriminatory subset using adaptive neural network ensemble consisting of three base classifiers such as Linear Vector Quantization (LVQ), Self-Organizing Map (SOM), Back Propagation algorithm (BP). Apply majority voting to combine probability distribution of base classifiers and evaluate predictive accuracy using test set / 10-fold cross validation method as performance evaluation measure.

3. BACKGROUND

3.1. Fuzzy Set Theory

Fuzzy set is an extension of conventional crisp set and was introduced by Zadeh (1965). A fuzzy set A is defined on a universe of discourse U, and is represented as a set of ordered pairs:

$$A = \left\{ \left(x; \mu_A \left(x \right) \right) \mid x \epsilon U \right\} \tag{1}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

where $\mu_A(x)$ is the membership value of $x \in U$ in fuzzy set A. The mapping $\mu_A: U \longrightarrow [0, 1]$ is established through membership function $\mu_A(.)$ to evaluate membership values in set A.

3.2. Rough Set Theory

This mathematical concept was introduced by Z. Pawlak (1982) to handle vagueness which is complementary to the concept of fuzzy set. Rough set approximates the conventional set by a pair of sets known as lower and upper approximation of the original set. The approximation / information space is denoted by a pair $<U, A>$, where U is the universe of discourse and A is the knowledge of U and the information function f is defined as: $f: U \times A \longrightarrow V$, where V is the domain of A. Therefore, an equivalence/ indiscernibility relation IND (P) can be defined on a subset P of the knowledge set A on U as:

$$IND\left(P\right) = \left\{\left(x_i, x_j\right) \epsilon U \times U \mid \forall a \epsilon P, f_a\left(x_i\right) = f_a\left(x_j\right)\right\} \tag{2}$$

$IND\ (P)$ creates a partition on U which is denoted as:

$$U \mid IND\left(P\right) = \left\{\left[x_i\right] P : x_i \epsilon U\right\} \tag{3}$$

where $[x_i]_P$ represents equivalence class or information granules. The elements of $[x_i]_P$ are equivalent or indiscernible with respect to the subset of attributes $P \in A$.

Any given arbitrary set $X \dot{\subseteq} U$, can be characterized in the approximation space $<U, A>$ by a pair of lower and upper approximations is shown in Figure 1 and defined as:

$$Lower - P\left(X\right) = U\left\{\left[x_i\right]_P \mid \left[x_i\right]_P \subseteq X\right\} \tag{4}$$

$$Upper - P\left(X\right) = U\left\{\left[x_i\right]_P \mid \left[x_i\right]_P \cap X \neq \Phi\right\} \tag{5}$$

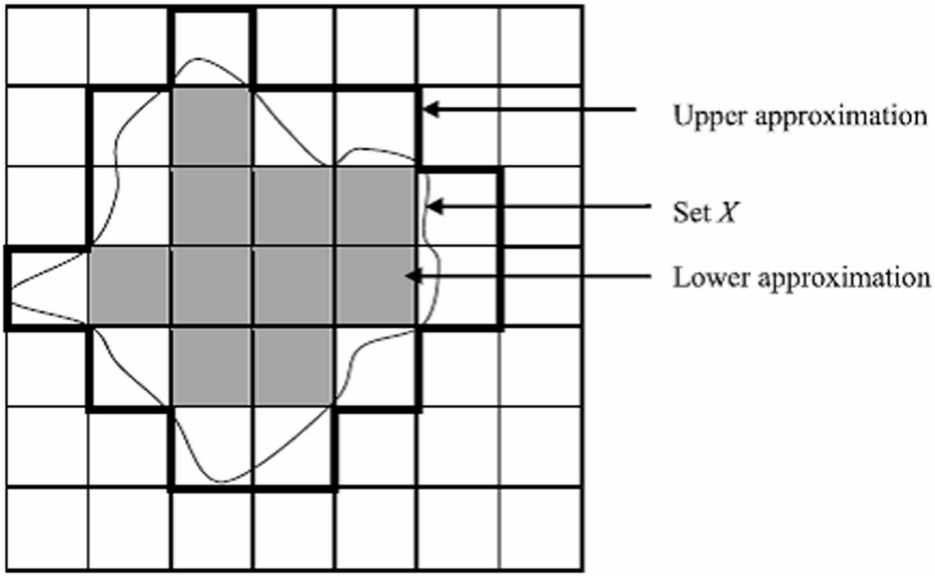
Now, set X can be represented as a pair of elements or tuple $<Lower\text{-}P(X), Upper\text{-}P(X)>$ in the approximation space $<U, A>$ and can referred to as rough set of X. Hence, $Lower\text{-}P(X)$ / $Upper\text{-}P(X)$ is the collection of those elements of U that belongs to X positively / possibly. The $Lower\text{-}P(X)$ is also called positive region and denoted as $POS_P(X)$.Set X is termed as exact set in $<U, A>$ if $Lower\text{-}P(X) = Upper\text{-}P(X)$ otherwise called a rough set. Boundary set is denoted by $BN_P(X) = Upper\text{-}P(X) \setminus Lower\text{-}P(X)$.

3.3. Fuzzy-Rough Sets

Rough sets generate a family of crisp equivalence classes by making crisp partitioning of the universe. Similarly, a family of fuzzy equivalence classes are generated by fuzzy equivalence relation or fuzzy knowledge granule (Jensen & Shen, 2004) by inducing a fuzzy partition on the universal set.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Lower and upper approximations of a rough set, X



Let represent the fuzzy equivalence space (Zhai, 2011) by <*U, A*> and a fuzzy subset of *U* by *X*. This indicates all decision and condition attributes are fuzzified. Then *fuzzy Lower-P and Upper-P* approximations can be defined as (Jensen & Shen, 2004):

$$Lower - \mu_{PX}\left(F_i\right) = \inf_x \left\{\max\left\{\left(1 - \mu_{FI}\left(x\right)\right).\mu_X\left(X\right)\right\}\right\}, \forall i \tag{6}$$

$$Upper - \mu_{PX}\left(F_i\right) = \sup_x \left\{\min\{\mu_{FI}\left(x\right)), \mu_x\left(X\right)\}\right\}, \forall i \tag{7}$$

where F_i is fuzzy equivalence class belongs to *U/P* and $\boldsymbol{\mu}_X(x)$ is the membership of x in X. However, the above definitions of fuzzy approximation deviate from the crisp upper and lower approximation, as the memberships are not available explicitly for all objects in the approximation space. Hence, fuzzy lower and upper approximations can be defined as (Xu, Miao & Wei, 2009):

$$Lower - \mu_{PX}\left(x\right) = \sup_{FI \in \frac{U}{P}} \min\left\{\mu_{FI}\left(x\right), Lower - \mu_{PX}\left(F_i\right)\right\} \tag{8}$$

$$Upper - \mu_{PX}\left(x\right) = \sup_{FI \in \frac{U}{P}} \min\left\{\mu_{FI}\left(x\right), Upper - \mu_{PX}\left(F_i\right)\right\} \tag{9}$$

Now the tuple <*Lower-P(X), Upper-P(X)*> is called a fuzzy-rough set which can be used for selecting real valued attributes from fuzzy lower approximation space.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.4. Genetic Algorithms

Genetic algorithms are evolutionary optimization methods motivated by biological phenomenon of natural selection and evolution and have wide applications in many different problems (Lee, Hong & Kim, 2009; Dash & Patra, 2013). It generates new population of chromosomes that is considered as the candidate solution in the iteration. Each chromosome is encoded as a string of binary, real, etc., version of a candidate solution. The fitness of each string of chromosome is then evaluated by a fitness function *f(x)*. Two fittest chromosomes having more fitness value are likely to be selected from the population and the genetic operator crossover and mutation are applied to generate new offspring from new population. The variation is introduced to the selected individuals for obtaining global optimum solution using crossover and mutation (Park & Cho, 2003; Luque-Baena, Urda, Subirats, Franco & Jerez, 2014).

4. THE PROPOSED LEARNING SCHEME

4.1. Evolutionary Hybrid Fuzzy-Rough-Genetic Search Method for Feature Selection

The first key element of the hybrid-FRGSNN model is its fuzzy-rough feature selection method which computes the reduct sets. Then an evolutionary search based wrapper model GSNN is employed on the outcome of filter FR method to select an optimal subset of features which is the outcome of the hybrid search FRGSNN model. Genetic algorithm is adopted here as a search tool to perform this operation. The reduct sets produced from fuzzy rough reduction process is used as the population of the genetic search which generates a reduced set of individuals in each new generation. In this scheme, *KNN* is used as a fitness function to evaluate the fitness of each individual elected from the population and then the operators like crossover and mutation are applied to reproduce new offspring. The primary motive of the proposed method is to optimize the search process to identify subset of features which improves the classification accuracy. It also focuses on reducing the size of the selected subset without compromising the accuracy rate of the problem. The detail configuration of genetic search is given below:

- **Encoding of Chromosome:** Each individual in the solution space is encoded in to a string of binary bits called chromosomes. Each bit of the string represents the state of each feature/gene in the reduct set derived from fuzzy-rough filter model (1 if feature is selected else 0).
- **Crossover:** A two-point crossover mechanism is employed to swap the substrings between the two points selected to generate new offspring.
- **Mutation:** Crossover cannot provide a satisfactory result if the encoded information of the population is not sufficient to solve the given problem. In such cases, a flip bit mutation operator is applied to the new offspring to generate a new chromosome spontaneously.
- **Selection:** Selection operator selects eligible parents for mating for the next generation. The parents are selected with a selection probability which is proportional to their fitness values. This operation is analogous to survival of the fittest in theory of evolution. In the proposed hybrid-FRGSNN model a binary tournament procedure will be used for selection of parents in each generation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4.2. Proposed Neural Network Ensemble for Classification

Adaptive neural network (Liu, Cui, Jiang & Ma, 2004; Luque-Baena, Urda, Subirats, Franco & Jerez, 2014) consists of a large number of massively interconnected neurons which function simultaneously to find solution of problems. Adaptive neural network (Lee, Hong & Kim, 2009) has the ability to learn how to perform tasks based on the data given for training or initial experience. Neural networks are efficient in recognizing patterns and classifying data in a large dataset with little information. They are best suited for predicting high dimensional datasets like microarray datasets. The architecture of the multilayer neural network consists of three types of layers of neurons such as input, hidden and output layer. The behaviour of the neural net increasingly depends on the connections between the components. These inputs and weighted connections between input and hidden layers determine the performance of the hidden layers. Similarly the performance of output layer is determined by the weight vector connecting hidden and output layer neurons. The objective of the network is to establish a mapping function which can minimize the error between desired and estimated output.

A single neural network cannot build a stable and generalized learning model for finding optimal solution for some specific problems. In this paper, we have ensemble (Yang P., Yang H., Zhou & Zomaya, 2010; Hansen & Salomon, 1990) three variants of neural networks such as LVQ, SOM and BP, in an attempt to find an efficient solution to this problem. Usually, the learners show some biasness to different attribute selection models (Zhou, Jiang, Yang & Chen, 2002). The performance of the ensemble depends on the meta-algorithm which aggregates the output of individual classifiers. A simple combination rule like majority voting has been applied in combining the resultant of the classifiers. The majority voting consensus for *k* classifiers in a given *n* classifiers where:

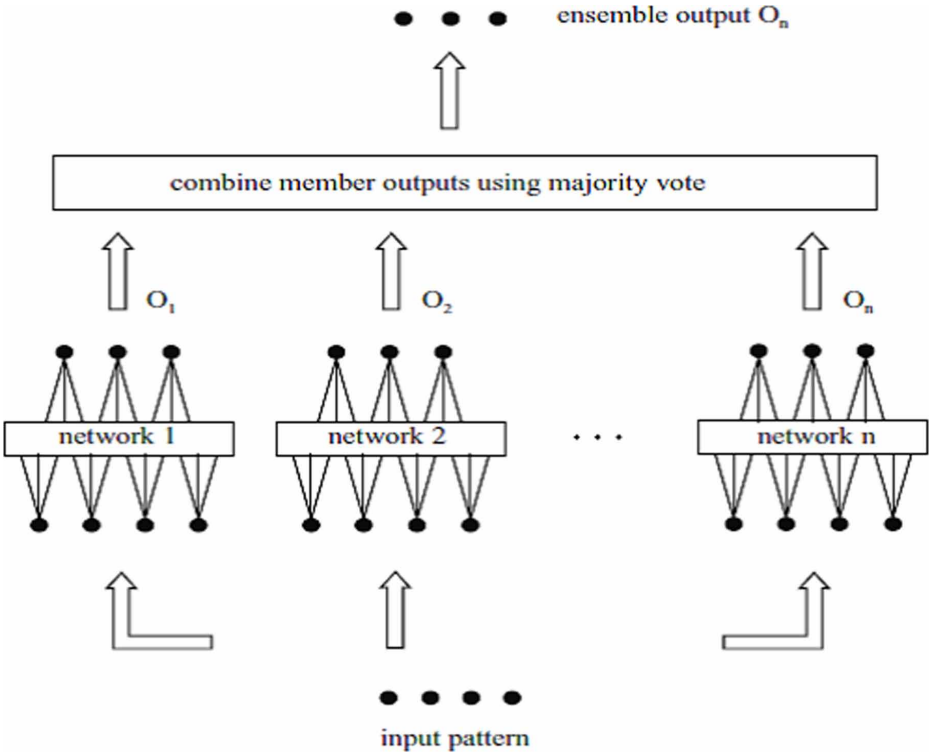
$$k = \begin{cases} \frac{n}{2} + 1 if\, n\, is evven \\ \frac{n+1}{2} if\, n\, is odd \end{cases}$$

The component classifiers used in the ensemble classifier namely LVQ, SOM and BP are combined and aggregated as a consensus committee. The subset of features produced by evolutionary fuzzy-rough feature selection phase will be fed into the ensemble classifier. The three base classifiers contained in this ensemble will use the input features to learn and classify the dataset separately. Finally, majority voting is used for calculating the consensus from the classification accuracy obtained from each classifier by using certain feature combination. Many significant advantages of ensemble learning algorithm over individual have been shown by many authors (Tan & Gilbert, 2003; Rosenwald, 2002).

The multi-net systems i.e., combination of a finite number of neural networks (Zhou, Jiang, Yang & Chen, 2002) has been used in different areas of research shown in Figure 2. The combined output of more than one classifier often provides improved accuracy over the output of individual network. However, the recent study of computational biology has witnessed an increasing use of ensemble learning because of the unique advantages in handling complex data structure, small sample size and high-dimensionality. The aim of this kind of ensemble learning is to achieve better generalization by combining the outputs of individual networks. The advantages of the ensemble learning i.e., how the problem of stability of the classifier can be improved compared to an individual classifier has been demonstrated in (Hansen & Salomon, 1990; Tan & Gilbert, 2003; Jiang, Yang & Chen, 2002).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Scheme of the ensemble



5. DATASETS ILLUSTRATION

A large amount of gene expression data is available in several datasets. In this paper, eight publicly available gene expression datasets are being used to evaluate the performance of the proposed feature selection and classification scheme. First five datasets, namely, DLBCL-NIH, DLBCL-Stanford, DLBCL-Harvard, CNS embryonal tumour outcome and Lung Harvard are collected from http://datam.i2r.a-star.edu.sg/datasets/krbd/ and last three, namely, ALL-AML, Brain Tumour and SRBCT are collected from http://www.gems-system.org/ which are illustrated in Table 1.

6. EXPERIMENTAL RESULTS AND DISCUSSION

In this paper, eight microarray datasets are used to ascertain the validity of the projected scheme. Out of these eight datasets, five datasets i.e., DLBCL–NIH, DLBCL-Stanford, DLBCL-Harvard, Central system (CNS) and ALL-AML are of binary category and the remaining three i.e., Lung-Harvard, Brain-Tumor-GEMS, SRBCT-GEMS belong to multi-class datasets. DLBCL data samples are taken from three different sources and studied their behaviour applying the proposed scheme. The databases utilized in this experiment are summarized in Table 1. All the learning algorithms used in this work are taken from WEKA machine learning package.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Gene expression microarray datasets

Datasets	Total number of genes	Training Samples	Test samples	References
DLBCL - NIH	7399	160(72:88)	80 (30:50)	Rosenwald et al., 2002
DLBCL - Stanford	4026	47 (24:23)	0	Alizadeh et al., 2000
DLBCL - Tumor	7129	77 (58:19)	0	Shipp et al., 2002
CNS	7129	60 (21:39)	0	Pomeroy et al., 2002
ALL_AML_GEMS	7129	38 (27:11)	34(20:14)	Golub et al., 1999
Lung Harvard	12600	122 (58:21:20:6:17)	81(81:0:0:0:0)	Bhattacharjee et al., 2001
Brain_Tumor1_GEMS	5920	90 (60:10:10:4:6)	0	Pomeroy et al., 2002
SRBCT_GEMS	2308	83 (29:11:18:25)	0	Khan et al., 2001

The experimental method is composed of feature selection and classification method. In feature selection process, the proposed hybrid FRGSNN model is applied to all datasets to identify subset of potential genes/ biomarkers by reducing the computational complexity of the problem. Traditional feature selection methods such as FR-filter and GSNN-wrapper model are applied to all datasets for comparison purpose. The parameters of genetic search algorithm used in wrapper and hybrid model are set as follows: crossover probability = 0.6, mutation probability = 0.033, maximum generation = 50 and population size = 100. The potential genes identified by hybrid FRGSNN, FR-filter and GSNN-wrapper are recorded in Table 3 for all the datasets. Then the representative subset of genes obtained from the above processes is classified by the proposed ensemble model and other learning algorithms. In addition to this, the reduced set of features are classified by classical ensemble methods namely, bagging, boosting and an advanced SVM ensemble. The performance of the learning algorithms is validated by using test dataset for DLBCL-NIH, ALL-AML and Lung cancer and ten-fold cross validation was applied to measure the statistically reliable predictive accuracy for the remaining datasets.

The proposed neural network ensemble is trained with three component neural networks and then combines the outputs of all the networks into an integrated output using majority voting (Liu, Cui, Jiang & Ma, 2004; Kuncheva, 2004). The subset of features is fed to each component classifiers, Bagging, Boosting ensembles and to the proposed ensemble for each dataset. Then the classification accuracy of each component classifiers and classical ensembles are compared with the accuracy obtained from proposed ensemble classifier. For all base neural network classifiers, the parameters are set as follows: BP learning is configured with 10 hidden nodes in a single hidden layer, output nodes equivalent to number of class of the datasets, learning rate varies from 0.01~0.10, momentum equals to 0.9 and maximum number of iterations fixed to 500. In LVQ learning, codebook vectors are fixed to 25, learning rate varies in the range 0.01~0.10, learning function is set to static, momentum sets to 0.1 and maximum iterations is 1000. In SOM learning, a 5 x 5 rectangle shape initial map is set with neighbourhood size 16, learning function sets to static, learning rate varies within the range 0.1~0.9 and maximum iterations is set to 2000. When they are used as independent classifiers, a small variation is made to the above parameters for learning purpose.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In this study, predictive accuracy *(A_{CC})* and kappa statistic *(k)* are used to measure the performance of learning algorithms. Predictive accuracy is computed using four base line statistical components (Liu, Cui, Jiang & Ma, 2004; Dash & Behera, 2016) namely, TP and FN denote the number of positive samples which are accurately and falsely predicted, respectively and TN and FP represent number of negative samples that are predicted accurately and wrongly, respectively:

$$A_{CC} = TP + TN / (TP + TN + FP + FN) \tag{10}$$

where TP denotes the correct classifications of positive samples, TN represents correct classifications of negative samples, FP represents the incorrect classifications of negative samples into positive class and FN is the incorrect classifications of positive samples into negative class. The description in Table 2 gives four baseline statistical components, where TP and FN denote the number of positive examples which are accurately and falsely predicted, respectively, and TN and FP represent the number of negative samples that are predicted accurately and wrongly, respectively.

Kappa statistic (*k),* as a statistical metric, measures the percentage of agreement (Kuncheva, 2004) between accuracy and diversity of base learners, which is shown along with the ensemble performance. The reliability of classification is measured using the formula:

$$k = \text{observed agreement} - \text{chance agreement} / (1 - \text{chance agreement}) \tag{11}$$

The value of *k* ranges from 0 to 1, when k= 1 shows a better agreement than value 0.

Table 2. Confusion matrix

	Predicted positive class	**Predicted negative class**
Real positive class	True positive (TP)	False negative (FN)
Real negative class	False positive (FP)	True negative (TN)

Table 3. Selection of genes using FR-filter, GSNN-wrapper and hybrid FRGSNN method

Datasets	**No. of genes before filtering**	**Selection of genes using FR-filter**	**Selection of genes using GSNN-wrapper**	**Selection of genes using hybrid FRGSNN**	**Percentage of genes selected by hybrid FRGSNN**
DLBCL-NIH	7399	294	3139	133	1.80
DLBCL-Stanford	4026	298	1205	124	3.08
DLBCL-Tumor	7129	162	2377	73	1.02
CNS	7129	162	2814	45	0.63
Lung Harvard	12600	234	3313	121	0.96
ALL-AML	7129	162	3442	58	0.81
Brain-Tumor	5920	267	2111	108	1.82
SRBCT-GEMS	2308	82	1041	20	0.87

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6.1. Results

We have aimed at finding a generalized prediction model for both binary and multi-class problems by exploiting the capability of ensemble by combining the output of base classifiers for predicting the solution of the problem. In this paper, we have compared the predictive accuracy of individual base learners as well as some established ensembles with the proposed ensemble neural network for all the available datasets using feature selection techniques.

Table 4 summarizes the classification accuracy of each component classifiers and proposed neural network ensemble on the subset of genes obtained from FR-filter, GSNN-wrapper and hybrid FRGSNN model. When results of the pair <FR-filter, classifier> and <GSNN-wrapper, classifier> compared with <hybrid-FRGSNN, classifier> methodology, it is observed that the pairing of proposed ensemble with hybrid-FRGSNN model yields better performance and also statistically significant difference is observed in all the datasets. It can also be observed from last column of Table 4 that hybrid FRGSNN model selects only 0.96% of genes from the largest Lung cancer dataset, 0.63% from CNS which is least and highest percentage 3.08% of genes from DLBCL-Stanford dataset.

The predictive accuracy of proposed ensemble algorithm for reduct set of genes obtained from FR-filter is better than the performance of component neural network learning algorithms. Similarly, for GSNN-wrapper model, comparison of results show a significant performance obtained by proposed ensemble method for six datasets achieving more than 90% accuracy. But, a highly remarkable performance is observed by ensemble method for the genes obtained from hybrid FRGSNN model for all the datasets in comparison to component classifiers. The analysis clearly indicates that the proposed learning scheme outperforms the predictive results of all three component classifiers. The same conclusion is drawn when compared with the findings of (Lee, Hong & Kim, 2009; Liu, Cui, Jiang & Ma, 2004; Tan & Gilbert, 2003).

Table 4 also provides a comparison between the classification accuracy of proposed ensemble classifier and the value of kappa statistic *(k)* for all the subsets. The classification accuracy and the value of kappa statistic *(k)* obtained from the proposed learning scheme has established a proper balance between accuracy and diversity. It is worth to note that, a small value of k denotes high diversity among the learners. Kappa statistic k measures the diversity between the pair of base learners, which is shown along with the ensemble performance. Hence, proposed method has achieved proper balance between diversity and accuracy and at the same time fulfilled the objective of finding a generalized prediction model.

Table 5 summarizes the predictive accuracy of the proposed ensemble and other ensembles such as B**agging and Boosting** for the subset of genes obtained from hybrid FRGSNN model for all the datasets. The analysis of the table proved that the other three ensembles work effectively for a particular microarray dataset, i.e., bagging performs well for ALL-AML, boosting does for Lung-Harvard using hybrid FRGSNN model. The predictive results obtained by Bagging and Boosting ensemble are not consistent across the datasets whereas the results obtained by neural network ensemble are remarkably significant across all the datasets.

Figure 3 shows the estimated accuracy obtained by all datasets using elementary classifiers such as LVQ, SOM, BP and the proposed neural network ensemble where Vote combines probability distribution of all three neural networks such as LVQ, SOM and BP using majority voting combination rule. We can see the accuracy obtained by ensemble is better for DLBCL-Harvard, ALL-AML, Lung-Harvard and Brain Tumor than other datasets. Here the performance of ensemble is enhanced by the hybrid FRGSNN

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Best estimated classification accuracy of three elementary neural networks and proposed ensemble method with estimated kappa statistic (k) value

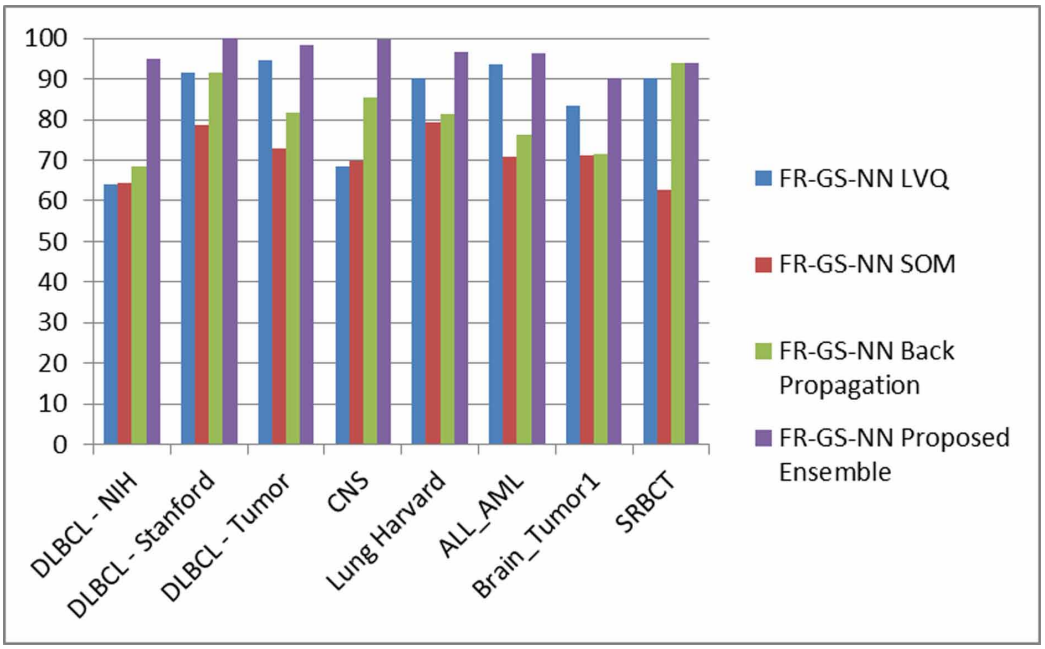
Datasets	FRGS				GSNN				FRGSNN			
	LVQ	**SOM**	**Back Propagation**	**Proposed Ensemble/ *k***	**LVQ**	**SOM**	**Back Propagation**	**Proposed Ensemble/ *k***	**LVQ**	**SOM**	**Back Propagation**	**Proposed Ensemble/ *k***
DLBCL - NIH	63.75	62.5	64.375	**93.125/0.86**	64.125	63.125	66.25	**90.625/0.81**	64.126	64.375	68.3333	**97.00/0.93**
DLBCL - Stanford	80.851	59.57	80.85	**100.00/1.00**	87.234	78.7234	89.3617	**100.00/1.00**	91.4894	78.7234	91.4894	**100.00/1.00**
DLBCL - Tumor	68.8312	57.1409	74.026	**84.4156/0.49**	92.2078	66.2338	81.8182	**94.8100/0.84**	94.5195	72.7273	81.8182	**98.5065/0.94**
CNS	68.33	66.6667	81.6667	**81.8200/0.41**	68.33	66.6667	84.33	**96.6700/0.92**	68.33	69.67	85.33	**99.6700/0.99**
Lung Harvard	86.6995	77.8325	68.4729	**89.1600/0.76**	89.6552	79.3842	80.3103	**92.6108/0.84**	90.1478	79.3842	81.4729	**100.00/1.00**
ALL_AML	73.6111	59.7222	69.4444	**80.5556/0.58**	90.8333	68.0556	70.8333	**94.4444/0.84**	93.4544	70.8333	76.3889	**99.4444/0.99**
Brain_Tumor1	74.444	66.6667	68.8889	**84.4444/0.72**	78.8889	68.3333	70	**87.7778/0.74**	83.3333	71.1111	71.6667	**90.2222/0.89**
SRBCT	74.6988	56.6265	87.9518	**90.3614/0.87**	86.747	62.6506	90.3976	**92.7229/0.84**	90.3614	62.6506	93.9759	**93.9759/0.91**

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Comparison of 10 fold cross validated predictive accuracy of proposed ensemble with other ensembles using features obtained from hybrid-FRGSNN

Ensembles Datasets	Bagging Accuracy (%)	Boosting Accuracy (%)	Proposed Ensemble Accuracy (%)
DLBCL - NIH	56.875	50.625	97.00
DLBCL - stanford	78.7234	78.7234	100.00
DLBCL - Tumor	77.9221	83.1169	98.5065
CNS	61.6667	58.3333	99.67
Lung Harvard	88.67	95.0739	100.00
ALL_AML	93.0556	87.5000	99.4444
Brain_Tumor1	78.8889	80.00	90.2222
SRBCT	83.1325	86.747	93.9759

Figure 3. Estimated accuracy of LVQ, SOM, BP and Neural Network Ensemble for all eight datasets in FR-GS-NN



feature selection method which selects optimal subset using genetic search. The accuracy of ensemble is ranges from 69.375% ~ 98.3333% with average of 83.854% whereas for element classifiers it ranges from 64.126% ~ 93.9759% with average of 79.44683%. This analysis clearly demonstrates, the proposed neural network ensemble outperforms the efficiency of element classifiers for all three types of feature selection methods (Lee, Hong & Kim, 2009; Liu, Cui, Jiang & Ma, 2004; Tan & Gilbert, 2003).

Table 6, shows the comparison of predictive results based on test set with the results published by Liu et al. (2004), Heesung Lee et al. (2009) and Tan and Gilbert (2003) for two major datasets. Again proposed learning scheme achieves 100% result which is significantly better than Tan and Gilbert (2003)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Comparison of predictive accuracy using test set

Author	Ensemble Classification	Datasets	
		ALL / AML	Lung
Tan and Gilbert (2003)	Bagged DT	91.18	93.29
Liu et al. (2004)	Neural Network Ensemble	100	100
Hessung Lee et al. (2009)	Neural Network Ensemble	98.82	-
Proposed Learning scheme	Neural Network Ensemble	99.444	100.00

Table 7. Comparison of predictive accuracy using 10-fold CV

Author	Ensemble Classification	Datasets	
		CNS	DLBCL
Tan and Gilbert, 2003	Bagged DT	88.33	-
Liu et al., 2004	Neural Network Ensemble	-	95.74
Proposed Learning Scheme	Neural Network Ensemble	99.67	100.00

for ALL-AML and Lung cancer in comparison to the results obtained by Liu et al (2004). In Table 7, using 10-fold cross validation, Tan and Gilbert (2003) reported 88.33% predictive accuracy for CNS and 95.74% for DLBCL reported by Liu et al. (2004) which is less than the proposed one.

6.2. Discussion

Feature selection plays an important role in enhancing the classification accuracy of the problem. Potential features influence the predictive ability of the classification model. In this study, a hybrid feature selection method and an ensemble neural network classifier are introduced for selection and classification of gene expression datasets. The proposed method outperforms the results published by other authors. We found the discrimination ability of the ensemble classifier is highly significant when search strategy like filter, wrapper and hybrid methods are concerned. Further, we used the same features for predicting component classifiers such as BP, SOM and LVQ and for two other ensembles such as bagging and Boosting and found that the results are not better than the proposed one. Hence, these significant improvements of the proposed scheme can be attributed to the fact that the evolutionary hybrid feature selection mechanism optimizes different combinations of genes that enhances the classification accuracy thereby improves the stability and generalization capability of the ensemble network.

The most important issue in ensemble learning is to establish a proper balance between the accuracy and diversity of the constituent base learners. In fact, the base learners used in an ensemble classifier need to have high classification accuracy and should avoid making misclassification errors. The diversity would be more among base learners if the misclassification error increases in the process of classification. It is also evident from Table 4 that the proposed methodology could provide a generalized prediction model which is the main objective of this research (Yang, Huei, Chuang & Yang C. Hong, 2009; kuncheva, 2004). In order to measure the level of agreement between a pair of base learners, a k-metric or kappa statistic (kuncheva, 2004) is used while correcting the chance.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

We believe that, ensemble classifiers are more appropriate for bioinformatics problems because it helps to design classifiers from noisy and incomplete data. Another inherent problem of microarray data is the relatively small number of samples compared to the dimension of the dataset. This problem is really a difficult task for a single learning classifier to induce an optimal hypothesis for this type of dataset. This can be appropriately addressed by ensemble learner which combines component classifiers. As a result, the network becomes stable and attains global minima instead of being trapped in local minima. But how many component classifiers need to be combined to capture the true hypothesis is still an open problem.

In this problem, we emphasized both feature selection and classification aspects of the problem. But we have not studied the importance of individual gene and relation among them which specifically help to get more information about the disease. This aspect of the problem is left for our future research.

7. CONCLUSION

Experimental results show that the evolutionary feature selection techniques effectively selected subset of marker genes for each component learner and increased the discriminatory power of the neural network ensemble. Also it is evident from the study that ensemble classifier performs better than classical ensembles and component classifiers for all datasets. The proposed method has achieved the highest averaged generalization ability compared to its counterparts and established an acceptable level of diversity among the base learners for majority of the analyzed benchmark datasets.

REFERENCES

Akadi, a. E., Amine, A., Ouardighi, A. E., & Aboutajdine, D. (2009). Feature selection for Genomic data by combining filter and wrapper approaches. INFOCMP Journal of computer science, 8(4), 28-36.

Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., ... Staudt, L. M. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, *403*(6769), 503–511. doi:10.1038/35000501 PMID:10676951

Bhattacharjee, A., Richards, W. G., Staunton, J., Li, C., Monti, S., Vasa, P., ... Meyerson, M. (2001). Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses, PNAS. *The National Academy of Sciences, USA*, *98*(24), 13790–13795. doi:10.1073/pnas.191502998 PMID:11707567

Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A., Benítez, J. M., & Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences*, *282*, 111–135. doi:10.1016/j.ins.2014.05.042

Chanho, P., & Cho, S. B. (2003. Evolutionary computation for optimal ensemble classifier in Lymphoma cancer classification, N. Zhonget al. (Eds.), Proceedings of ISMIS ‘03, LNAI (Vol. 2871, pp. 521–530).

Chen, T., Hong, Z., Deng, F., Yang, X., Wei, J. & Cui, M. (2015). A novel selective ensemble classification of microarray data based on teaching-learning-based optimization. International journal of multimedia and ubiquitous engineering, 10(6), 203-218.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., ... & Meltzer, P. S. (2001). Classification and diagnostic of cancers using gene expression profiling and artificial neural networks. *NCBI*, *7*(6), 673–679. PMID:11385503

Dash, S. (2015). A Rule Induction Model Empowered by Fuzzy-Rough Particle Swarm Optimization Algorithm for Classification of Microarray Dataset. In Computational Intelligence in Data Mining, SI*ST* (Vol. 3, pp. 291-303). doi:10.1007/978-81-322-2202-6_26

Dash, S. (2016). Hybrid Ensemble Learning Methods for Classification of Microarray Data, Book Chapter published in Handbook of research on Computational Intelligence Applications in Bioinformatics (pp. 17–36). Hershey, PA, USA: IGI Global. Doi:10.4018/978-1-5225-0427-6

Dash, S., & Behera, R. (2016). Sampling based Hybrid Algorithms for Imbalanced Data Classification. *International Journal of Hybrid-Intelligent-Systems*, *13*(2), 77–86. doi:10.3233/HIS-160226

Dash, S., & Patra, B. N. (2012). Rough Set Aided Gene Selection for Cancer Classification. *Proceedings of 7th International Conference on Computing and Convergence Technology (ICCCT)* (pp. 290-294).

Dash, S. & Patra, B.N. (2012). A Hybrid Data Mining Technique for Improving the Classification Accuracy of Microarray Data Set. I.J. Information Engineering and Electronic Business, 4(2), 43-50.

Dash, S., & Patra, B. N. (2013). Redundant Gene Selection based on Genetic and Quick-Reduct Algorithms. *International Journal on Data Mining and Intelligent Information Technology Application*, *3*(2), 1–9.

Diaz-Uriarte, R., & de Andreas, S. A. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*(1), 3. doi:10.1186/1471-2105-7-3 PMID:16398926

Ghosh, A., & Jain, L. C. (Eds.). (2005). *Evolutionary computation in data mining*. Berlin: Springer-Verlag. doi:10.1007/3-540-32358-9

Golub, T. R., & ... . (1999). Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring, 1. *Science*, *286*(5439), 531–537. doi:10.1126cience.286.5439.531 PMID:10521349

Guyon, I., Gunn, S., Nikravesh, M., & Zadeh, L. A. (Eds.). (2006). *Feature Extraction: foundations and applications*. Berlin: Springer. doi:10.1007/978-3-540-35488-8

Hansen, L., & Salomon, P. (1990). Neural networks ensemble. *Pattern Analysis and Machine Intelligence*, *12*(10), 993–1001. doi:10.1109/34.58871

Huang, F., Zhou, Z., Zhang, H., & Chen, T. (2000). Pose invariant face recognition. *Proceedings of 4th IEEE International Conference. Automatic Face and Gesture Recognition* (pp. 245–250).

Jensen, R., & Shen, Q. (2004). Fuzzy–rough attributes reduction with application to web categorization. *Fuzzy Sets and Systems*, *141*(3), 469–485. doi:10.1016/S0165-0114(03)00021-6

Kuncheva, L. (2004). *Combining Pattern Classifiers*. John Wiley &Sons. doi:10.1002/0471660264

Lee, H., Hong, S., & Kim, E. (2009). A new genetic feature selection with neural network ensemble. *International Journal of Computer Mathematics*, *86*(7), 1105–1117. doi:10.1080/00207160701724760

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lee, H., Hong, S., & Kim, E. (2009). Neural Network Ensemble with probabilistic fusion and its application to gait recognition. *Neurocomputing*, *72*(7-9), 1557–1564. doi:10.1016/j.neucom.2008.09.009

Liu, B., Cui, Q., Jiang, T., & Ma, S. (2004). A combinational feature selection and ensemble neural network method for classification of gene expression data. *BMC Bioinform.*, *5*(1), 136–147. doi:10.1186/1471-2105-5-136 PMID:15450124

Luque-Baena, R. M., Urda, D., Subirats, J. I., Franco, L., & Jerez, J. M. (2014). Application of genetic algorithms and constructive neural networks for the analysis of microarray cancer data. *Theoretical Biology & Medical Modelling*, *11*(Suppl. 1), S7. doi:10.1186/1742-4682-11-S1-S7 PMID:25077572

Pawlak, Z. (1982). Rough sets. *Int. J. Inform. Comput. Sci.*, *11*(5), 341–356. doi:10.1007/BF01001956

Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., ... Golub, T. R. (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, *415*(6870), 436–442. doi:10.1038/415436a PMID:11807556

Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I., ... Staudt, L. M. (2002). The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *The New England Journal of Medicine*, *346*(25), 1937–1947. doi:10.1056/NEJMoa012914 PMID:12075054

Saeys, Y., Inza, I., & Larranaga, P. (2007). A review of feature selection techniques in Bioinformatics. *Bioinformatics (Oxford, England)*, *23*(19), 2507–2517. doi:10.1093/bioinformatics/btm344 PMID:17720704

Shipp, M. A., & ... . (2002). The AP1-dependent secretion of gelectin-1 by Reed-Sternberg cells fosters immune privilege in classical Hodgkin lymphoma, PNAS. *The National Academy of Sciences, USA*, *104*(32), 13134–13139.

Tan, A. C., & Gilbert, D. (2003). Ensemble machine learning on gene expression data for cancer classification. *Applied Bioinformatics*, *2*, S75–S83. PMID:15130820

Witten, I. H., Frank, E., & Hall, M. A. (2011). Data mining: practical machine learning tools and techniques (3rd ed.). Morgan Kaufmann series in data management systems. Menlo Park: Morgan Kaufmann.

Xu, F.F., Miao, D.Q., & Wei, L., (2009). Fuzzy-rough attributes reduction via mutual information with an application to cancer classification. Computers and Mathematics with Applications, 57(6), 1010-1017.

Yang, C. Huei, Chuang, L. Y., & Yang, C. Hong. (2009). IG-GA: A hybrid Filter/Wrapper Method for Feature Selection of Microarray Data. *Journal of Medical and Biology Engineering*, *30*(1), 23–28.

Yang, P., Yang, Y. H., Zhou, B. B., & Zomaya, A. Y. (2010). A review of ensemble methods in bioinformatics. *Bioinformatics (Oxford, England)*, *5*(4), 296–308.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. doi:10.1016/S0019-9958(65)90241-X

Zhai, J. (2011). Fuzzy decision tree based on fuzzy-rough technique. *Soft Computing*, *15*(6), 1087–1096. doi:10.100700500-010-0584-0

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zhou, Z.-H., Jiang, Y., Yang, Y.-B., & Chen, S.-F. (2002). Lung cancer cell identification based on artificial neural network ensembles. *Artificial Intelligence in Medicine*, *24*(1), 25–36. doi:10.1016/S0933-3657(01)00094-X PMID:11779683

KEY TERMS AND DEFINITIONS

Ensemble Method: Ensemble methods are learning algorithms that construct a set of classifiers and then classify new data points by taking a weighted vote of their predictions. It's of two types: homogeneous and heterogeneous ensemble.

Fuzzy-Rough Feature Selection (FRFS): Provides a means by which discrete or real-valued noisy data or a mixture of both can be effectively reduced without the need for user-supplied information.

Genetic Algorithm: Genetic algorithms are evolutionary optimization methods motivated by biological phenomenon of natural selection and evolution.

Hybrid Ensemble Learning: The potential of ensemble learning is augmented by combining the characteristics of another learning algorithm

Neural-net Ensemble: Neural network ensemble can significantly improve generalization accuracy of networks by training several networks and combining their results.

This research was previously published in the International Journal of Knowledge Discovery in Bioinformatics (IJKDB), 6(1); edited by Mehdi Khosrow-Pour, D.B.A.; pages 1-16; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 37
Wired and Wireless Distributed e-Home Healthcare System

Booma Devi Sekar
University of Macau, Macau

JiaLi Ma
University of Macau, Macau

MingChui Dong
University of Macau, Macau

ABSTRACT

The proactive development in electronic health (e-health) has introduced seemingly endless number of applications such as telemedicine, electronic records, healthcare score cards, healthcare monitoring etc. Yet, these applications confront the key challenges of network dependence and medical personnel necessity, which hinders the development of universality of e-health services. To mitigate such key challenges, this chapter presents a versatile wired and wireless distributed e-home healthcare system. By exploiting the benefit of body sensor network and information communication technology, the dedicated system model methodically integrates some of the comprehensive functions such as pervasive health monitoring, remote healthcare data access, point-of-care signal interpretation and diagnosis, disease-driven uplink update and synchronization (UUS) scheme and emergency management to design a complete and independent e-home healthcare system.

INTRODUCTION

Along with socioeconomic progress and development, three global crises come into being that threatens human medical service industry, including population aging, prevalence of chronic diseases, and soaring health cost. Driven by falling fertility rate and increasing life expectancy, older population (above age 65) is escalating with surprising speed and will outnumber young children (under age 4) in the very near future. As reported by World Health Organization (WHO, 2011), between 2010 and 2050, the propor-

DOI: 10.4018/978-1-7998-1204-3.ch037

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tion of older adults will double from 7.7% to 15.6%. While global population aging represents a triumph of medical, social, and economic advances over diseases, it also presents tremendous challenges to the society and healthcare services. What's more, such demographic changes result in inevitable increase in the prevalence of chronic diseases or non-communicable diseases (NCDs). As the leading cause of fatality throughout the world, according to Mathers and Loncar (2006), NCDs take up 63% among the total deaths and the aged-related NCDs mortality is projected to rise to 69% in 2030. Among the four main types of NCDs, cardiovascular disease (CVD) is prevalent and closely linked with aging population. According to the survey conducted by WHO in 2008, 17 million, which represents 30% of global deaths, were due to CVD and it is estimated that by 2030, this could further increase to 23 million. To effectively prevent the mortality rate due to CVD, early detection and timely treatment would be the ideal choice. However, this imposes huge burden on the medical personnel during long-term monitoring and diagnosis. To make things worse, the cost of healthcare service is increasing continuously and outpaces economic growth in gross domestic product (GDP). As revealed by the third national health survey by Liu et al. (2007), during 1998–2003 the annual increase of healthcare expense was 11.8% whilst the average income of people in rural areas increased by only 2.4% annually during the same period. The skyrocketing health cost not only prevents lot of people from seeking for early medical care, but also brings about large disparities in service quality and accessibility among different income groups and between rural and urban populations (Powell, 2005).

Therefore, in view of demographic and epidemiologic needs and escalating expenditure of the current healthcare system, improving healthcare service is a top priority issue throughout the entire world. Targeting at developing long-term, low-cost, sustainable, and accessible healthcare service, practical enactment of taking preventive measures in the traditional healthcare environment has always been a challenge. Thus, exploiting the advancement of multiple key technologies including information & communication technology (ICT), mobile technology, wireless body sensor networks (WBSN), intelligent technology etc., research over the past decade has focused on improving the quality, safety and efficiency of healthcare services by creating an electronic healthcare (*e*-healthcare) environment. With the aim to provide healthcare to anyone, anytime and anywhere, the development of *e*-healthcare has not only improved the convenience and quality of healthcare services towards patients and healthcare consumers, but also has drastically facilitated medical personnel to pursue early detection and diagnosis of various pathologies. In realizing the vision of *e*-healthcare systems, there exist numerous technical and non-technical challenges. Here, some of the crucial requirements and challenges are listed as follows.

1. **Privacy and Security:** Since wireless/wired transfer of biomedical signals and pathological information are concerned in *e*-healthcare systems, user privacy should be guaranteed to avoid malicious and illegal abuse. Security of the entire system, WBSN, and wireless/wired network is of great significance (Kara, 2001; Zhang et al., 2002);
2. **Reliability:** To ensure the successful message transfer between healthcare consumers and providers, reliability of the networks should be strictly controlled with low error tolerance. Particularly for wireless network, robust infrastructure and protocols are demanded to support the reliability and quality of wireless communication (Varshney, 2007);

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3. **Patient Engagement:** Taking into consideration of various users with varying education level/ knowledge, an accessible *e*-healthcare system with user-friendly operations and comprehensible functions are desired. Especially for elder population who have very little interests in accessing state-of-the-art technologies, the key challenge allies on effectively engaging them to properly use such systems (White et al., 2006);
4. **Interoperability:** With the advent of more and more *e*-healthcare systems, interoperability may become a serious issue, such as semantic intelligibility (Adibi, 2012), interoperability between different platforms, interference with medical devices (Varshney, 2007), etc.
5. **Agreement and Other Management Issues:** Such as ownership of medical and health data, multilayer infrastructure management (Adibi, 2012), payment (Varshney, 2007), etc.

During the past decades, many *e*-healthcare systems have been proposed (Adibi, 2012; Caldeira et al., 2012; Kyriacou et al., 2003; Niyato et al., 2009; Patra et al., 2009; Ren et al., 2010; Varshney, 2007; Wang et al., 2010; White et al., 2006;). Such developments include remote health monitoring (Varshney, 2007; Ren et al., 2010; Niyato et al., 2009; Wang et al., 2010), telemedicine (Kyriacou et al., 2003), electronic health records (EHR) (Patra et al., 2009), intelligent emergency management system, medical information management system etc. Some of the successful models from literature include a remote patient monitoring system by integrating multiple wireless technologies (Niyato et al., 2009). In this model, before delivering the acquired signals to healthcare service provider, heart rate is first recorded and its severity is analyzed. If the heart rate is examined to be in the critical range, the medical record of that particular patient is transmitted to the hospital with higher priority by temporarily bypassing all the other patients' data in queue. Based on telecommunication technology, a telemedicine system is proposed in (Kyriacou et al., 2003). Such a system consists basically of two units: telemedicine unit at hospital site for performing effective emergency management and a base unit at patient's location for receiving and displaying the incoming data to the patient. Also to support the functional operation of telemedicine, literature show that researchers have explored various key technologies to enable remote health record access, a distributed EHR system (Patra et al., 2009) could be cited as a particular example. Finally, with the trend of developing mobile health (m-health) system on mobile devices, Adibi (2012) presents an m-health solution on BlackBerry mobile phones. It uses mobile networks (3G/4G), mobile phones work as a transmission and communication platform for end-to-end data transmission.

Yet, it is evidenced that though such developments offer various benefits for improving healthcare services, they suffer from some critical challenges. First of all, such systems are incapable of self-diagnosis and leave the burden of interpretation to hospital doctors. This in fact increases the burden on medical personnel as they are now required to attend to both on-site and remotely connected patients. Secondly, these systems are network dependent and rely on cable or wireless network to transmit signals and receive results from hospital server or centralized computation station, thus such models are not suitable for rural areas or network unavailable users.

In light of above, this chapter focuses on transforming the hospital-rescue-centered healthcare scheme to patient-prevention-centered approach and introduces a framework of wired and wireless distributed *e*-home healthcare system for CVD health monitoring. The dedicated system principally integrates the comprehensive functions, such as pervasive health status monitoring, remote healthcare data access, point-of-care signal interpretation and diagnosis, disease-driven uplink update and synchronization (UUS) scheme and as well as emergency management to support various users diversified medical demands.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The functional architecture of the proposed model consists of three functional modules, namely upstream (hospital server), midstream (patient's local computer) and downstream (embedded-link mobile devices). In the downstream, the biomedical signals, specifically electrocardiogram (ECG), sphygmogram (SPG) and heart sound (HS) are acquired from the subject with the proposed multi-vital sign acquisition system. To efficiently work with the resource limitation of portable mobile devices in downstream, a preliminary health monitoring model is designed to monitor the cardiac condition through analysis of hemodynamic parameters derived from SPG signal. Through exploiting the advancement in AI technology, an efficient and more reliable intelligent monitoring and diagnosis model is built on the local computer in midstream. The proposed model in midstream executes a hierarchical analysis and interpretation of various biological signals to provide comparatively more reliable point-of-care to the user. Finally, as both down- and mid-stream form as independent health monitoring modules implemented in the portable mobile devices and local computers for home users respectively, it becomes critical to introduce some element of intervention during emergency and necessity situations. Thus, upstream module which is installed in the hospital server is developed. The upstream module is mainly designed with a top-level intelligent CVD monitoring and diagnosis system (TICMD). As the users of *e*-home healthcare system ultimately rely on the diagnosis made by upstream module during emergency and necessity situations, the proposed TICMD system employs both non-invasive biological signals such as SPG, ECG, HS etc., and invasive blood test parameters (BTP) to diagnose various CVD. With the support of a stronger knowledge base from doctors and data base from acquired patients' records, the diagnosis results are highly accurate. Through implementing system-level hierarchical diagnosis, both on-line and locally off-line diagnosis with ranked accuracy is delivered to users, thus satisfying various user demands in varying situations.

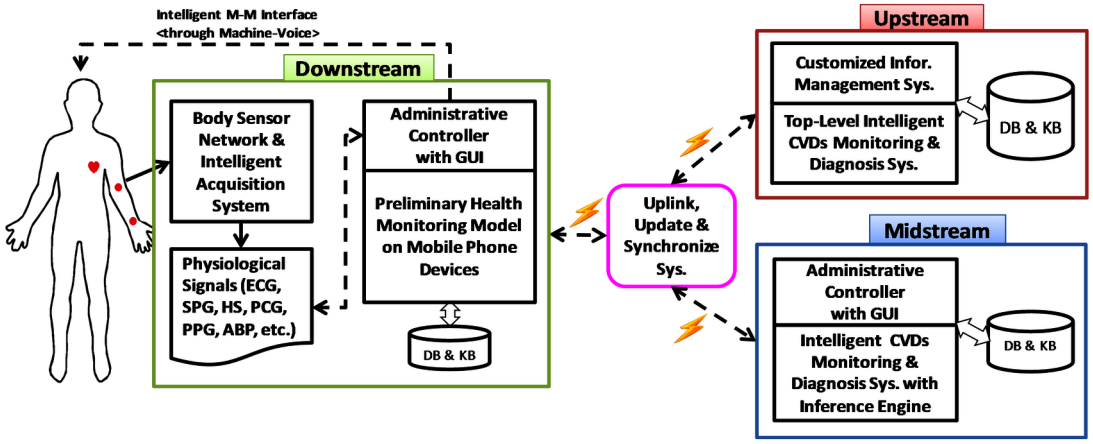
Finally, to systematically interconnect the three modules of *e*-home healthcare system, an UUS scheme is proposed. The UUS acts as a specialized backyard on-line assistance to support the communication between thousands of home users, healthcare center and hospital server. Thus, allowing the distributed *e*-home healthcare system to work as a whole. The main benefit of the proposed *e*-home healthcare system is that at normal conditions it would function independently and only at emergency situations it would automatically connect with hospital server for further support.

FUNCTIONAL ARCHITECTURE OF DISTRIBUTED *e*-HOME HEALTHCARE SYSTEM

The functional architecture of distributed *e*-home healthcare system for CVD health monitoring is illustrated in Figure 1. The proposed architecture can be majorly divided into three functional modules, which include the downstream (embedded-link devices e.g. smart phone, iPad, PDA etc.), midstream (local computer) and upstream (hospital server). Furthermore, these modules are interconnected by a mutual network communication through our proposed UUS system. The hierarchical ordering of the modules also represents the hierarchical diagnosis process, in the sense that the downstream with its limited computational ability, storage space and knowledge base (KB) serves with a quick preliminary diagnosis results. Subsequently the midstream, which consists of an intelligent diagnosis model, designed using various signal processing techniques and hybrid artificial intelligence (AI) technologies provides a better diagnosis results in a home healthcare environment. Finally, the upstream which incorporates KB from hospital with the information acquired from mid-stream and down-stream, provides a high-precision diagnosis results to the patient.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Overall functional diagram of the wired and wireless distributed e-home healthcare system for CVD diagnosis



Now, to decide on which of the above module would be sufficient or necessary to make an efficient and reliable diagnosis decision, the UUS is introduced. UUS basically acts as the backend support and remote assistance system which interconnects the down- and mid- stream with the up-stream of the system model. Whenever the diagnosis system in lower streams encounters doubt in the results or failed diagnosis caused by insufficient KB, the UUS is triggered to send help request to the remotely-linked upstream. UUS principally manages the requests, feedback the absent KB rule to lower streams and perform re-diagnosis with the help of customized information management system built using abundant KB rules and comprehensive data base (DB) of all user clients' information. Overall, the key significance of UUS is that it covers the deficiency and limitation in lower streams and supports various user demands, thereby effectively supports the functionality of the overall distributed *e*-home healthcare system.

In the following sections, to understand the operation of the proposed wired and wireless distributed *e*-home healthcare system, in this chapter the key features and operation of each functional module will be further elaborated in details.

Downstream Functional Module

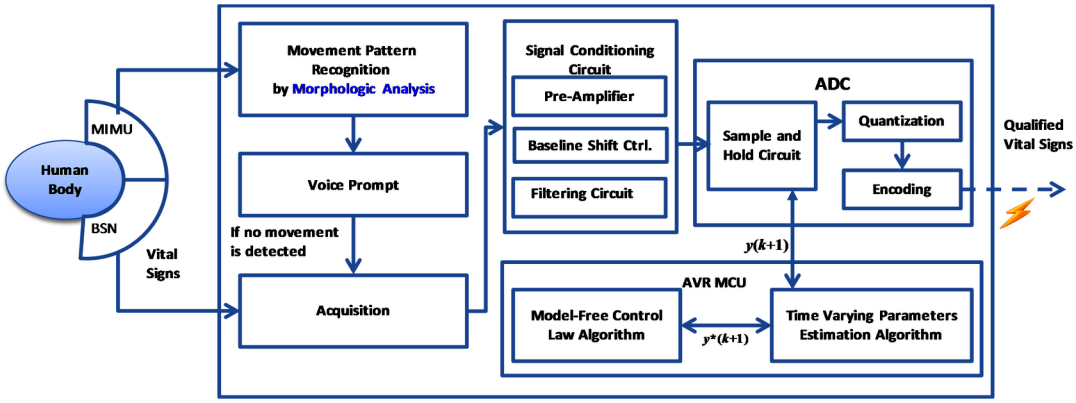
The downstream mainly comprises of an intelligent portable/wearable biological signal acquisition system and a preliminary health monitoring model designed based on WBSN and embedded link system. The main functionality of the downstream is to efficiently acquire and store multi-vital signs including ECG, SPG and HS from the subject and provide preliminary health monitoring status of the patient.

Intelligent Wearable Biological Signal Acquisition System

The rapid advancement in wireless communication and network technologies, linked with the progress of multi-sensor data fusion and miniaturized pervasive health monitoring devices, led to the evolution of WBSN and embedded link system technologies. To facilitate the research and development (R&D) of pervasive health monitoring system, such key technologies are exploited to design the architecture of an intelligent portable/wearable multi-vital sign acquisition system.

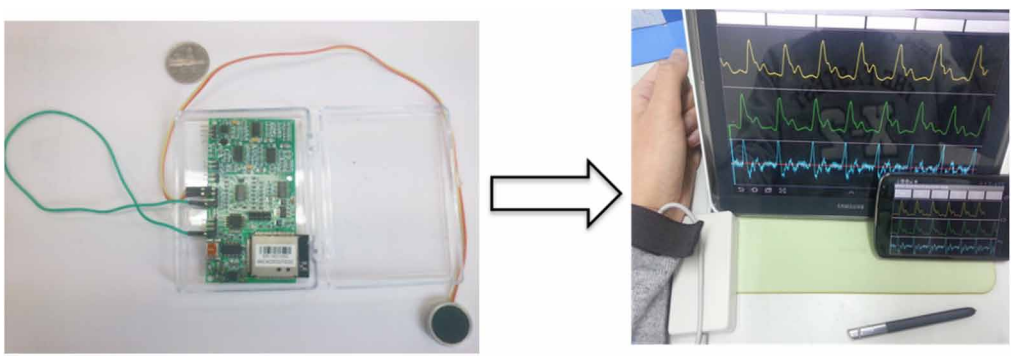
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Architecture of an intelligent portable/wearable multi-vital sign acquisition system



The designed multi-vital sign acquisition system presented in Figure 2 comprises of biosensors, micro inertial measurement unit (MIMU), signal pre-processing and conditioning module and a closed-loop control unit. The biosensors attached to the user through embedded link system detect multi-vital sign, whereas the MIMU detects any physical movement hindering the stable biological signal acquisition process. Through the proposed spatial movement tracking and pattern recognition approach (Gan et al., 2014), MIMU recognizes the movement pattern of the user and immediately initiates a voice prompt to guide the user in proper operation of the biological signal acquisition system. Once the MIMU detects stabilized movement pattern, the bio-sensors attached to the user through embedded link system would automatically sample and record the biological signals. The sampled signals would then be subsequently processed in the signal pre-processing and conditioning module. Here, the sampled signals are de-noised through filtering and normalized with pre-amplification and baseline shift. In addition, a closed-loop control unit with an analog to digital (ADC) converter is constructed with automatic voltage regulator (AVR) and micro control unit (MCU) to adaptively self-adjust and record stable multi-vital sign. A prototyping system designed with the proposed architecture is presented in Figure 3. Here, the designed embedded-link device transmits the acquired multi-vital sign to the portable device in the downstream and to the patient's local computer in the midstream through wireless/wired transmission for further analysis, interpretation and diagnosis.

Figure 3. Architecture of an intelligent portable/wearable multi-vital sign acquisition system

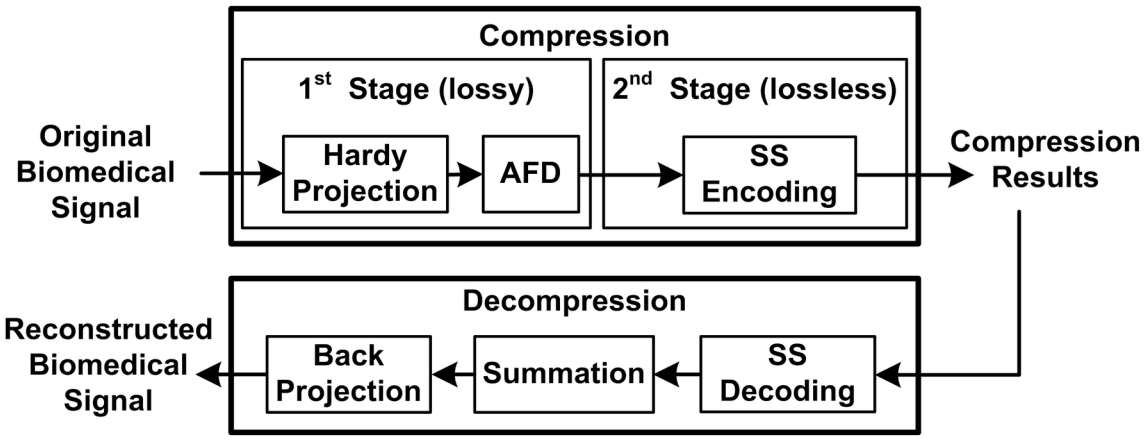


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

One aspect worth noting here is that as the proposed model is designed to be linked with portable devices such as smart phone, iPad, PDA etc., minimum storage space becomes the main limitation. However, massive multi-media medical data generated during long-term monitoring and diagnosis not only imposes a huge threat to such source-limited devices in terms of efficient storage but also introduces unaffordable burden during band-restricted mutual communication. To address such key challenges, an effective and reliable data compression technique to efficiently store and retrieve biomedical signal is proposed by hybridizing adaptive Fourier decomposition (AFD) and symbol substitution (SS) technique (Ma et al., 2015). The functional design of proposed method for compressing and decompressing biomedical signals is depicted in Figure 4.

In the proposed model, for biomedical signal compression, two stages are considered. First, after projecting the input signals into hardy space, AFD performs lossy compression by decomposing projection signals into weighted polynomial basis components in modified Takenaka-Malmquist (TM) system (Qian et al., 2011). Different from other transform compression methods, AFD adaptively generates tailor-made basis according to various input signals, implying comprehensive applicability for compressing multi-vital signs. With the aid of maximal projection principle (MPP), only a few items are enough to reconstruct the input signals with high efficiency and fidelity. In the second stage, AFD decomposition results are further compressed losslessly using SS technique. SS substitutes the compressed results of long bits with sequence number of short bits referring to a customized encoding table. In the meanwhile, SS fulfills built-in data encryption simultaneously benefiting from shuffling encoding tables. Therefore it alleviates the challenge of information disclosure in most existing methods. As for decompression, it is the inverse process of compression and possesses quite low computation complexity. After decoding the SS substitution results, the projection signal is obtained by summing up the weighted components. Finally, the original input signal is recovered with little distortion through back projection. By validating using ECG, SPG, and HS signals, the proposed method exhibits a good tradeoff between compression efficiency and reconstruction quality, pushing forward the compression performance to an unexploited region.

Figure 4. Compression and decompression method using AFD and SS technique



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Preliminary Health Monitoring Model for Portable Device

Once the multi-vital signs are acquired by the embedded-link device, the immediate expectation of the user is to obtain the diagnosis result in his/her portable device. Thus, in the proposed design, taking into consideration of the resource limitation of portable device to store and process adequate knowledgebase (KB) and inference rules, a preliminary health monitoring model to efficiently provide warning messages to patient is developed.

The design of proposed model consists of three important parts, which include the KB framework design, KB customization and finally the inference engine to analyze the cardiac condition and generate warning messages to the user. Now, as knowledge management in resource-limited downstream unit is quite challenging, only the KB and inference rules generated with 32 hemodynamic parameters (HDP) derived from SPG signal (Li et al., 2005) are stored and analyzed in portable mobile device connected with the embedded-link system. The other biological signals including ECG and HS are analyzed by the patient's local computer in midstream.

For developing the proposed preliminary health monitoring model for portable device, Java development kit with its common adaptability on embedded-link devices is chosen. However it must be emphasized that the proposed model can also be developed on other software platform.

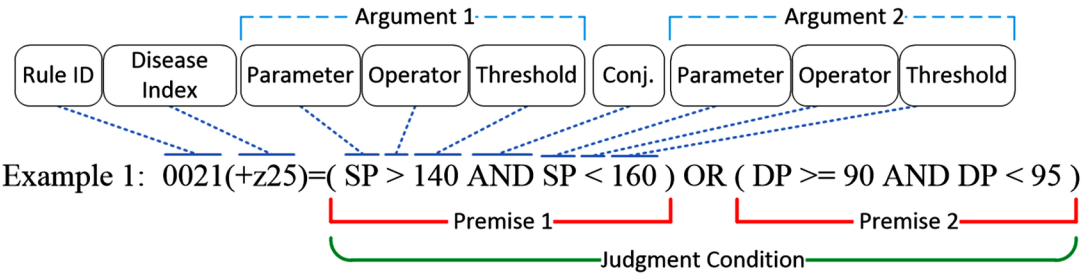
Knowledgebase Framework Design and Customization

In the design of KB framework for a potable mobile device, some of the important technical specifications considered necessary include

1. Storage of the essential medical data and inference rules in limited resource;
2. Define the inference rules in a systematic and consistent structure;
3. Extensibility to perform efficient knowledge update from the server;
4. Efficiently retrieve specific knowledge for the diagnosed pathology.

Thus, in the proposed KB framework, with the acquired KB from doctors, the inference rules are predefined in machine language and stored in text format for easy access and process (Choi et al., 2014). As illustrated in Figure 5 the structure of inference rule is defined in a systematic format, which composes

Figure 5. Inference rules for cardiovascular disease knowledgebase



Example 2: 0010(+z11)=(VPE >= {1.2*[2*Weight+45]*0.0112})

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of three parts: rule identification (ID) number, disease index and judgment condition. Here, each rule is identified by its rule ID number, which could range from 0001 to 9999. Due to resource limitation and also as only preliminary diagnosis is performed in portable device; proposed model consists of inference rules generated through the analysis of HDP derived from SPG. In a specific inference rule, the disease index is represented by "z" followed by a unique number, which corresponds to a specific pathological condition. Here, the symbol "+/-" indicates that the succeeding judgment condition is for/against the listed disease. As for the judgment condition, multiple premises can be represented in the same rule by defining them within brackets and distinguishing them with a conjunction/disjunction node. Finally, argument is the basic component of a premises, which compares the real value of certain biomedical parameter, computed from the site-sampled user's vital sign with a pre-defined fixed threshold value or a condition presented within "{ }". Two inference rules defined using the proposed KB framework is shown in Figure 5. The main advantage of the proposed KB framework is that it is resource saving as each rule with varying complexity would occupy approximately 60 ~ 300 bytes, which implies that on an average 1000 bytes of storage space can accommodate more than 1000 rules, an absolute necessary feature to be considered in the design of KB framework for portable device. Furthermore, the inference rules are stored in text format, the proposed KB framework can be applied in various mobile devices and as well as in the local computer in midstream. Also, as the inference rules are defined in systematic format, it becomes relatively easier for performing KB update and synchronization between local computer in midstream and the portable device in downstream.

Knowledgebase Customization

As in any diagnostic model, the reliability and accuracy of the model will directly rely on the diversity of rules in local KB framework. However, it is difficult to store all the useful rules in resource-limited portable mobile devices. Therefore, in the proposed approach a customized KB is generated to utilize the storage space in local computer in midstream and offset the resource limitation in portable device.

In the customized KB model, the pre-defined inference rules are divided into two groups: fundamental disease group and specific disease group. The fundamental disease group will basically contain all the inference rules relevant to the diagnosis of common types of CVD, while the specific disease group would contain sub-groups of inference rules for diagnosing different disease categories. In the proposed model, the KB framework in the portable mobile device only contain the fundamental disease group and the specific disease group being larger in size would be stored in the local computer memory. However, to increase the reliability of diagnosis in portable device, at the startup of *e*-home healthcare system, the user will be allowed to further customize the portable device by downloading one or a set of specific disease group. For example, if the user is suffering from hypertension and needs to specifically monitor the relevant parameters regularly, he/she could download the sub-group of inference rules relevant to hypertension into the portable mobile device. This would allow the user to not only monitor his cardiac system as a whole but also specifically monitor his/her hypertension levels, reliably at the convenience of the portable mobile device.

Inference Engine

With the constructed KB framework, the proposed inference engine will adapt the basic inference tree structure to perform preliminary diagnosis on the embedded-link portable platform. As the inference rules in individual branches of the tree are pre-defined, the inference engine will hierarchically judge

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the conditions of the relevant branch based on the acquired biomedical parameters and subsequently diagnose the potential heart pathology. In the diagnosis process, the inference engine will record the corresponding disease indexes whose inference rules have been satisfied. Then based on the inference results, the user will be indicated on his/her health status, possible/potential heart pathologies and the relevant warning messages. The warning message is an important feature of the preliminary CVD diagnosis model in a portable device as it could guide the user to take timely action. Basically, the warning message could include one or more of the following similar messages:

1. Improve the cardiac function with some exercise;
2. Pay attention to a particular biomedical parameter;
3. Update the KB by downloading some specific disease group package;
4. Test your biomedical signals in local computer;
5. Seek medical treatment, etc.

In conclusion, the preliminary health monitoring model on portable device provides the user with comprehensive health status information to help the user to regularly monitor his/her health status and take appropriate action accordingly.

Testing Results of Downstream Module

To validate the intelligent wearable biological signal acquisition system and the preliminary health monitoring model proposed and designed for the downstream module, the following tests are conducted.

First of all, as the downstream module is operated by the user in a home environment, it is highly critical to guide the user in recording stable multi-vital sings for further analysis. Thus, as proposed a MIMU is mounted on the biosensor (e.g. SPG acquisition unit placed on the wrist) to detect any movement during recording. Based on spatial movement tracking, the MIMU basically detects any movement of the wrist or forearm during SPG signal recording. To test the MIMU, a volunteer is requested to make some specific movements during recording to observe the corresponding trajectory results. The trajectory results in Figure 6(a) and 6(b) show that the MIMU could very well track the rectangular and triangular movement prompted by the volunteer. This indicates that the MIMU could effectively recognize any type of unexpected movement from the user during recording biological signal and accordingly guide in proper signal acquisition through a voice prompt.

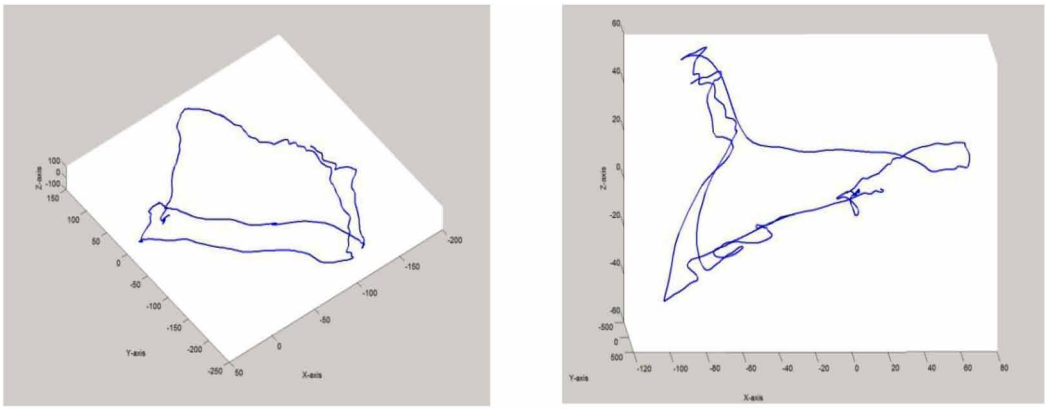
For validating the data compression technique to efficiently store and retrieve multi-vital sign, three ECG records acquired from lead I from MIT-BIH arrhythmia benchmark database (Mark & Moody, 1998), two SPG records acquired from patients from hospital in China and two HS records acquired from eGeneralMedical.com (n.d.), USA, database are utilized.

Three indices namely compression ratio (CR), percentage root mean square difference (PRD) and the normalized PRD (PRDN) as defined in Equations 1, 2, and 3 are assessed to measure the quality of compression output of the biological signals.

$$CR = \frac{b_o}{b_r} \tag{1}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. (a) Rectangular spatial tracking result; (b) triangular spatial tracking result



(a) (b)

where, b_o and b_r denote the number of bits in original and reconstructed signals respectively.

$$PRD\left(\%\right) = \sqrt{\frac{\sum\left(f-g\right)^2}{\sum f^2}} \tag{2}$$

$$PRDN\left(\%\right) = \sqrt{\frac{\left(f-g\right)^2}{\left(f-f_m\right)^2}} \tag{3}$$

where, f, g represents the original and recovered signals respectively and f_m is the mean value of original signal.

In assessing the compression results, higher CR and lower PRD and PRDN are expected to attain an efficient and reliable data compression. Table 1 compares the testing results of the proposed method with the ones proposed in literature.

It is observed from Table 1 that the proposed method exhibits competitive performance with existing state-of-the-art works presented in literature.

Finally, to examine on whether the downstream module could operate as an independent unit, a patient suffering from hypertension tests the unit by recording his biological signals and examining the parameters in the mobile portable device connected to the *e*-home healthcare system. The result displayed in the mobile device is illustrated in Figure 7.

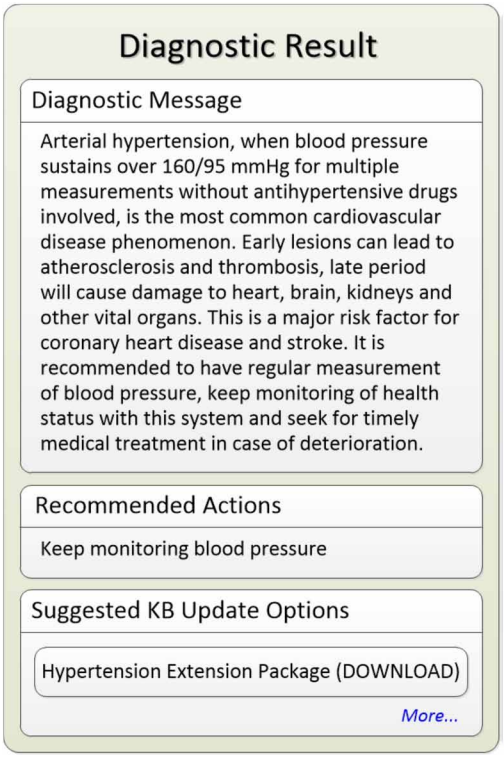
The warning message in the mobile device indicates the patient to monitor the blood pressure regularly. Also, the hypertension extension package is listed in the interface for customized update. In conclusion, the preliminary health monitoring model in the downstream module provides the user with comprehensive health status information to help users to monitor health status and take appropriate actions accordingly.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Validation results of proposed data compression approach

Method	Performance Indices	ECG			SPG		HS	
		Rec 1	Rec 2	Rec 3	Healthy	Patho.	Healthy	Patho.
Proposed model	CR	14.19	20.41	20.21	10.11	7.94	65.26	60.08
	PRD	0.32	0.48	0.77	1.25	1.42	1.94	1.98
	PRDN	9.22	8.92	8.80	3.20	4.44	1.94	1.98
DCT (Lee et al., 2011)	CR	22.94	24.43	19.31	N/A	N/A	N/A	N/A
	PRD	1.95	1.17	2.05				
	PRDN	48.88	20.80	16.30				
HermiteTrans. (Ma et al., 2013)	CR	14.24	17.06	20.22	10.31	11.23	N/A	N/A
	PRD	0.73	0.81	0.61	1.07	1.19		
	PRDN	20.48	15.03	3.72	2.69	3.25		
Wavelet Trans. (Manikandan & Dandapat, 2007)	CR	N/A	8.39	N/A	N/A	N/A	52.65	
	PRD		4.10				4.66	

Figure 7. A diagnostic result of a hypertension patient shown on a mobile device



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MIDSTREAM FUNCTIONAL MODULE

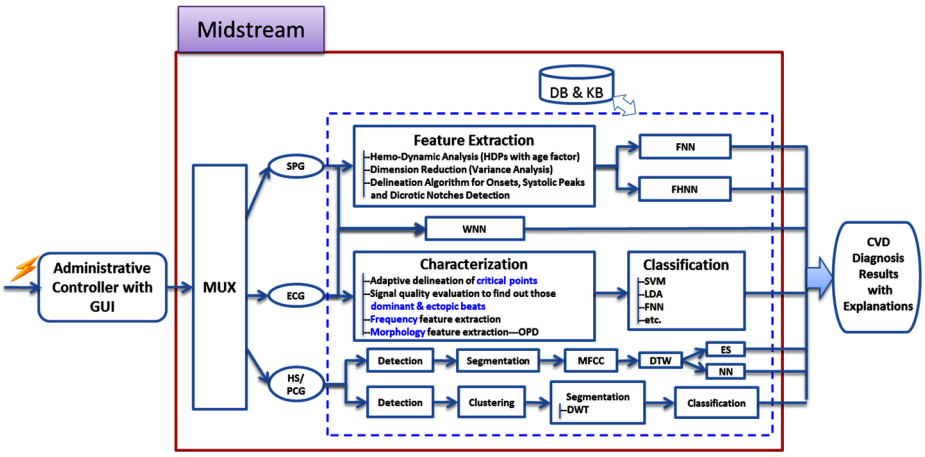
As the main goal of the proposed wired and wireless distributed *e*-home healthcare system is to transform hospital-rescue-centered healthcare to patient-prevention-centered approach, it is crucial that an efficient and reliable intelligent monitoring and diagnosis system (IMDS) is built on the local computer in midstream. Thus, by exploiting the advancement in AI technology, the proposed IMDS executes a hierarchical analysis and interpretation of various biological signals and provides a more efficient and reliable point-of-care CVD diagnosis to the user. The proposed architecture of the IMDS is illustrated in Figure 8. Here, the pre-processed multi-vital signs are transmitted to the local computer through wired or wirelessly connected embedded-link biological signal acquisition system. In the midstream, as all the three biological signals including SPG, ECG and HS/PCG are analyzed, the reliability and accuracy of the IMDS is comparatively more powerful than the preliminary health monitoring model built on the portable device in downstream.

In the proposed architecture of IMDS, a user-friendly administrative graphic user interface (GUI) is provided to help the user to conveniently interact with IMDS in biological signal analysis and diagnosis. Based on the diagnosis output, the user will be prompted to take some precaution or connect to the remotely-linked diagnosis system in the upstream for an even more accurate diagnosis using both non-invasive and invasive parameters. Yet, as the main goal of the proposed *e*-home healthcare system is to limit the need of connecting to the hospital server in the upstream, our research team have researched and developed various algorithms and techniques to diagnose CVD using the non-invasive biological signals including SPG, ECG and HS/PCG signals (Choi et al., 2014; Lei et al., 2007a, 2007b; Sekar et al., 2011, 2012; Shi et al., 2009; Zheng et al., 2013). In the following sections, some of the more efficient and reliable methods developed will be described in details.

Intelligent Diagnosis System using SPG

SPG is one of the most fundamental biomedical signals and carries profuse medical information needed for health monitoring of cardiovascular system. SPG signal can provide clinically useful information beyond the brachial measure of blood pressure and could ideally reflect four different aspect of a car-

Figure 8. Architecture of intelligent monitoring and diagnosis system in midstream



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

diovascular system, including the heart functionality (e.g. coronary artery status, oxygen consumption), arterial status (e.g. elasticity, peripheral resistance and compliance), blood status (e.g. blood viscosity and volume) and microcirculation function (e.g. circulation time period).

From literature study, one could note that majority of research have been devoted towards the development of various medical devices/instruments for recording arterial blood pressure (ABP) and SPG signal (Hwang et al., 2004; Im & Lessard, 1995; Miao, 1992). Whereas, another stream of research have been devoted to defining cardiovascular system model to evaluate the cardiac function and vascular properties of the cardiovascular system. One such study is the derivation of hemodynamic parameters (HDP) from SPG signal for analyzing the pathological condition of heart (Ursino et al., 1996). This requires extraction of morphological feature from SPG signal and derivation of HDP thereof. These HDP would include series of critical risk factors, such as cardiac output (CO), stroke volume (SV), total peripheral resistance (TPR) systematic arterial compliance (SAC) etc., derived from the morphological features of the SPG signal. The proposed diagnosis model in midstream mainly focuses on analyzing these HDP for CVD classification.

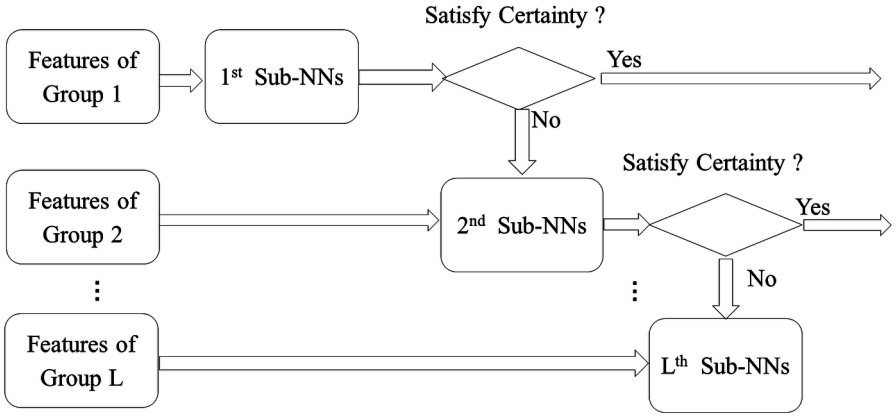
In the following sections, an intelligent diagnosis model designed based on fused hierarchical neural networks (FHNN) will be elaborated in details.

Fused Hierarchical Neural Networks

With advancement in AI technology, recent studies show that integration of AI technologies such as fuzzy logic (FL), artificial neural networks (ANN), expert system (ES) etc., have their unique advantage in scientific research. In this study, integration of ANN and hierarchical networks is proposed to design a layered neuronal structure to more effectively solve a diagnosis/classification problem. Based on the exploration of some of the hierarchical formation (Bagher-Ebadian et al., 2004; Calado & Sa da Costa, 1998; Kovalerchuk et al., 1997) from literature, a fused hierarchical neural networks as shown in Figure 9 is proposed.

According to the proposed scheme, FHNNs are hierarchically constructed by using diverse sub-NNs. The input features are categorized into a number of groups based on their influence on discriminating the hypotheses, group 1 with highest importance/priority and so on. Then as shown in Figure 9, the 1st

Figure 9. Proposed scheme of fused hierarchical neural networks



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

sub-NNs would intake features of group 1, and will attempt to deduce conclusion rapidly. If outputs of 1st sub-NNs satisfy the pre-set acceptance criteria, that is if the deduction certainty is satisfied, the outputs of 1stsub-NNs would yield as the overall output result of FHNNs; else the 2nd sub-NNs would be introduced, which would intake features of group 2 along with the midway outputs of former sub-NNs as inputs. Sequentially, the features, group by group from high to low priority, are in taken by subsequent sub-NNs; until either the deduction certainty is satisfied or all features are completely used.

With such scheme introduced, the question now arises to see how such a scheme can be adapted for a particular application. Specifically, the key challenges now would be to categorize the input parameters according to their importance/priority and then propose appropriate sub-NNs in each layer of the hierarchical structure to efficiently solve the classification problem. Thus, in this chapter with a systematic approach, methods to address both these key challenges are proposed and an intelligent CVD diagnosis model using HDP is designed and validated.

Categorization of Input Parameters

To design the proposed model, site-sampled medical data are acquired from two hospitals in China, namely Beijing Changping Chinese Medicine hospital and the first affiliated hospital of Guangxi Medical University. These include each patient's physiological attributes, original SPG waveforms, HDP and doctor's clinical diagnostic results. Here, the medical symptom space is denoted as $MSS \in R^N$, where $N = 38$, includes 6 physiological attributes and 32 HDP. Totally, 2267 medical records of 165 patients were randomly collected.

These samples were first pre-processed to remove any outliers and then with the implementation of variance analysis (Gelman, 2005), *F*-statistic is computed to categorize the 32 HDP and 6 physiological parameters according to their importance/priority in classifying three types of CVD, including coronary heart disease (CHD), hypertension (HT), and hyperlipidemia (HL). Table 2 shows the categorization results obtained using variance analysis. Based on the confidence coefficients range of [99%, 100%], [95%, 99%] and [0, 95%] respectively, 4 parameters were classified into the sensitive (more important) group, 12 parameters as the auxiliary and the remaining 22 parameters as the replenish (less important) group.

Selection of Sub-NNs for FHNN

According to the proposed scheme of FHNN presented in Figure 9, one of the main key challenges was to identify the most suitable type of NNs structure in each layer of FHNN that would be efficient in solving the classification problem. Thus, in the proposed approach to support such selection, the categorized input parameters in Table 2 were tested on three different NNs structure, including pure fuzzy NNs (FNNs), fuzzy probabilistic NNs (FPNNs) and fuzzy learning vector quantization networks (FLVQ) in diagnosing three types of CVD (CHD, HT and HL) and their combination. The structural architecture of these NNs will be discussed later. Now, such testing results presented in Table 3, clearly shows which category of input parameters are best analyzed with what type of NNs structure.

According to Table 3, FPNNs and FLVQ obviously surpass the basic FNNs in classifying the three types of CVD and their combination. Now in comparing FPNNs and FLVQ, FPNNs have better performance in utilizing sensitive and auxiliary HDP groups to detect CVD, whereas FLVQ performs better in utilizing replenish HDP group. This inspires one to construct FHNNs by using FPNNs as 1st as well as 2nd sub-NNs and FLVQ networks as the 3rd sub-NNs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Input parameters grouped through variance analysis

Sensitive	Auxiliary	Replenish	
• Age • CO • FEK • PAP	• Weight • SP • DP • PR • SI • CI • HOV • CMBV • PP • MSP • SPR • yr	• SV • VPE • EWK • SWI • HOI • CMBR • MDP • MAP • CCP • BLK • AC • TPR	• VER • PAWP • BC • y • MHR • MRT • MST • Height • Sex • PAR

Acronyms are clarified in Table 1 in (Shi et al., 2010).

Table 3. FNNs, FPNNs, and FLVQ exploration on categorized input parameters

Method	HDP Groups Elected as Input	Accuracy in Detecting CHD	Accuracy in Detecting HT	Accuracy in Detecting HL	Accuracy in Detecting Mixed CVD
FNNs	Sensitive	51.46	76.20	40.18	34.54
	Auxiliary	50.09	76.46	20.89	29.50
	Replenish	52.50	77.85	31.32	34.69
	Use all features at once	72.69	75.00	49.54	34.72
FPNNs	Sensitive	59.72	88.89	70.83	47.92
	Auxiliary	71.76	75.46	71.30	45.37
	Replenish	70.83	77.78	68.52	48.61
	Use all features at once	76.39	68.98	60.65	46.30
FLVQ	Sensitive	67.36	85.53	79.86	45.83
	Auxiliary	73.38	78.59	68.87	44.21
	Replenish	72.92	93.17	70.83	54.63
	Use all features at once	80.65	92.87	68.89	52.78

Network Structure of Sub-NNs

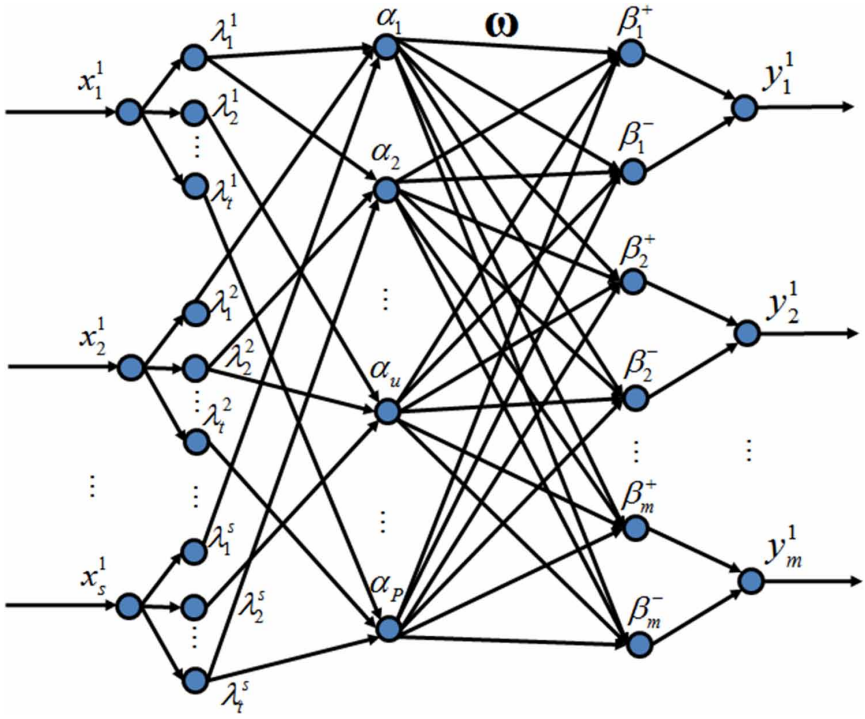
1. Fuzzy Probabilistic Neural Networks

The probabilistic neural networks (PNNs) would basically execute a Bayesian decision making process. PNNs generally memorize the training cases, so during testing if it is presented with a similar case; the PNNs would normally classify it accurately. Literature show, PNNs and FPNNs have been applied for solving various classification problems (Asgary & Mohammadi, 2005; Georgiou & Vrahatis, 2008; Mishra et al., 2008; Song et al., 2007).

The proposed FPNNs shown in Figure 10 are a five layer network structure, comprising of input, fuzzification, pattern matching, synthesis and output layers. As FPNNs is used as the sub-NNs in the first and second layers of FHNNs in Figure 9, the number of inputs in the first and second sub-NNs

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Network structure of FPNNs



equals to the number of parameters in the sensitive and auxiliary group respectively. Moreover, thesecond sub-NNs will also analyze the output of third layer of first sub-NNs for further classification. Here, a detailed exposition of the FPNNs as executed in first layer of FHNNs will be made.

In Figure 10, the input vector of the 1st layer of FPNNs is denoted as $x = \left[x_1^1, x_2^1, \ldots, x_s^1\right]^T$. In the 2nd layer, the input vectors are fuzzified to obtain the corresponding fuzzy membership grade (FMG) with $\lambda \in [0,1]$. The membership functions are pre-defined based on the age factor adjustment as described in our paper (Shi et al., 2010). Denoting the level of fuzzification as *t*, the number of output vectors from 2nd layer would be $t \times s$.

The 3rd layer computes the matching degrees α to fuzzified training patterns using Equation 4.

$$\alpha_u = \frac{1}{s\sigma} \sum_{i-1}^{s} e^{-\frac{(\lambda_i - \lambda_i^u)^2}{\sigma^2}} \tag{4}$$

where *u* is the u^{th} number of nodes in the 3rd layer, its maximum value equals to the number of training patterns *P*; λ_i is the FMG of i^{th} sensitive feature; and $\sigma$is smoothing parameter.

The 4th layer consists of 2*m* nodes to reflect two types of output: *m* positive and *m* negative CVD nodes. The nodes summate the corresponding training pattern matching degrees together according to Equation 5, and then forward it to the 5th layer.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\beta_j^+ = \sum_{j=1}^{P}\left(\omega_j^+ \cdot \alpha_j\right), \beta_j^- = \sum_{j=1}^{P}\left(\omega_j^- \cdot \alpha_j\right) \tag{5}$$

where ω_j^+ is j^{th} link weight connecting j^{th} positive CVD nodes, ω_j^- caters for negative CVD in similar way. ω in Figure 10 denotes link-weight matrix comprising of β_j^+ and β_j^-, j=1, 2, …, P.

The 5th layer with m category nodes is the output layer, which selects the sign of winner between positive and negative CVD assumed nodes as output according to Equation 6:

$$y_k^1 = \left(\beta_k^+ - \beta_k^-\right), \left(k = 1,2,...,m\right) \tag{6}$$

where y_k^1 represents the k^{th} CVD output of the 1st sub-NNs, positive y_k^1 means k^{th} CVD patient and negative y_k^1 means that it is free of k^{th} CVD. Absolute value of $\left(\beta_k^+ - \beta_k^-\right)$ is defined as the certainty for diagnosing k^{th} CVD.

2. Fuzzy Learning Vector Quantization Networks

Learning vector quantization (LVQ) networks is a special type of ANNs, in the sense that LVQ classification algorithms do not approximate density functions of class samples and directly define class boundaries based on the prototype, a nearest-neighbor rule and a winner-takes all paradigm. Therefore, it maintains rapid convergence and minimizes classification errors. FLVQ has been specifically applied for solving pattern recognition problem (Baghshah et al., 2006; Jatmiko et al., 2009).

In the proposed model, FLVQ is used as the sub-NNs in the 3rd layer of FHNNs to analyze the input parameters from the replenish group. The input vector is thus denoted as $\boldsymbol{x}^{\boldsymbol{L}} = \left[x_1^L, x_2^L, ..., x_r^L\right]^T$, here L=3.

The input vectors are then fuzzified as FMG and represented as vector $\boldsymbol{\kappa}$ in the $2n^d$ layer of FLVQ. When the fuzzification level remains t, then the number of output in $2n^d$ layer is computed as t×r.

$$\boldsymbol{\kappa} = \left[\kappa_1^1, \kappa_2^1, ..., \kappa_t^1, \kappa_1^2, \kappa_2^2, ..., \kappa_t^2, ..., \kappa_1^r, \kappa_2^r, ..., \kappa_t^r\right]^T \tag{7}$$

The 3rd layer of FLVQ produces a vector that represents the negative of distances between FMG and vectors formed by link-weights. The v^{th} row of link-weight matrix $\boldsymbol{\eta}$ can be denoted as

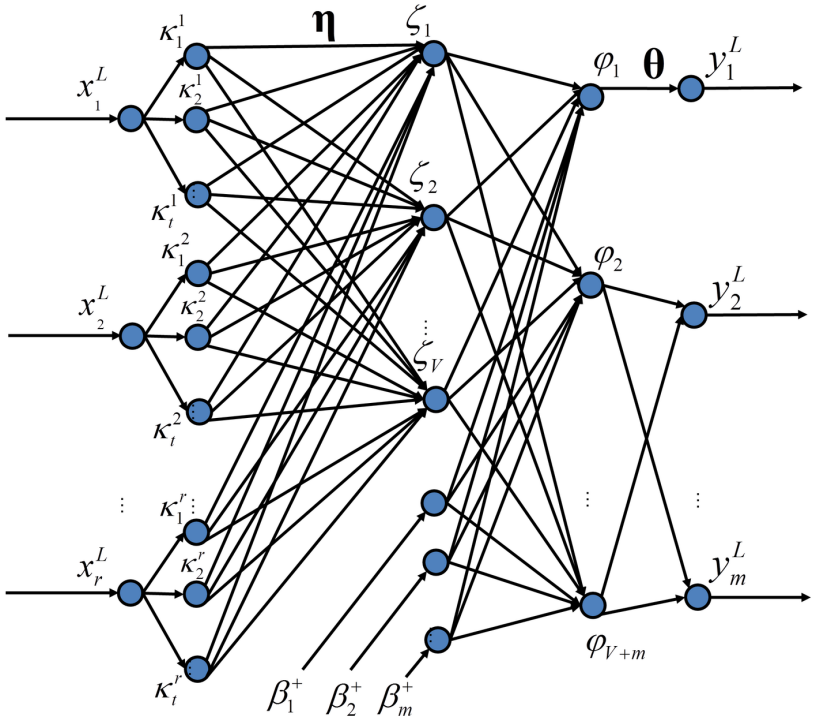
$$\boldsymbol{\eta}^{\boldsymbol{v}} = \left[\eta_1^v, \eta_2^v, \ldots, \eta_{t\times r}^v\right].$$

Here, the maximum v is denoted as V, which is the number of learning patterns as well as the number of nodes in 3rd layer. The output of 3rd layer is marked as

$$\boldsymbol{\zeta} = \left[\zeta_1, \zeta_2, ..., \zeta_V\right]^T = \boldsymbol{\eta} \times \boldsymbol{\kappa},$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Network structure of FLVQ



which quantizes the similarity between FMG and learning vector.

The 4th layer outputs of former 2nd sub-NNs are presented as partial input to the 3rd layer of FLVQ. The outputs β^+ of *m* positive CVD nodes are taken as inputs to next FLVQ layer. So there are *V*+*m* nodes in 3rd layer of FLVQ.

The 4th layer is a competitive layer, composed by same number of nodes as 3rd layer. The competitive layer accepts inputs similarly and returns neuron outputs of zero for all neurons except for the winner. The output vector is denoted as ϕ, it's *i*th element is calculated by following Equations 8, 9, and 10.

$$\varepsilon = \max\left(\zeta_1, \zeta_2, ..., \zeta_V, \beta_1^+, \beta_2^+, ..., \beta_m^+\right) \tag{8}$$

$$\phi_i = \begin{cases} 1, & \forall \zeta_i = \varepsilon \\ 0, & \forall \zeta_i < \varepsilon \end{cases}, when\ i = 1, 2, ..., V \tag{9}$$

$$\phi_i = \begin{cases} 1, & \forall \beta_{i-V} = \varepsilon \\ 0, & \forall \beta_{i-V} < \varepsilon \end{cases}, when\ i = V+1, V+2, ...V+m \tag{10}$$

The 5th layer transforms the output of competitive layer into predefined target classifications. The patterns learned by the competitive layer are referred to as subclasses of categories.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The outputs of the 3rd sub-NNs, FLVQ, are calculated by $\left[y_1^L, y_2^L, ..., y_m^L\right]^T = \boldsymbol{\theta} \times \boldsymbol{\varphi}$, where L=3, $\boldsymbol{\theta}$ is link-weight matrix of the last layer with m rows by $V+m$ columns.

The purpose of training FLVQ is to adjust link-weight matrix $\boldsymbol{\theta}$. For every input vector, the networks will give a CVD classification result. If the result of classification is correct, the connection weights value can be corrected by Equation 11:

$$\theta_k\left(\tau+1\right) = \theta_k\left(\tau\right) + \xi * \left(y_k^L - \theta_k\left(\tau+1\right)\phi_\tau\right) \tag{11}$$

If it is wrong, the link-weights should be adjusted by Equation 12:

$$\theta_k\left(\tau+1\right) = \theta_k\left(\tau\right) - \xi * \left(y_k^L - \theta_k\left(\tau+1\right)\phi_\tau\right) \tag{12}$$

where $\xi \in (0,1)$, $\theta_k(\tau)$ is τ^{th} training vector of k^{th} category, $k = 1, 2, \ldots, m$.

This process is repeated until the achieved classification rate is satisfied or the maximum epoch number is reached. After learning, the FLVQ is ready and the whole FHNNs can be used to detect the CVD.

Intelligent Diagnosis System Using ECG

ECG is perhaps the most commonly recognized and used biomedical signal so far. In general practice, cardiologists manually examine the tracing of ECG recording to identify any abnormalities. As such procedure is oppressive and time-consuming, literature show that abundance research has focused on developing automated methods of ECG interpretation and classification. Some of the related researches include ECG signal de-noising (Oster et al., 2010), detection-delineation algorithms to detect critical points (Ghaffari et al., 2009), signal stability and quality evaluation (Zhang et al., 2011), feature point characterization (Husser et al., 2009), recognition and classification to assist medical diagnosis (Chiang et al., 2010).

It can be noted from literature study that the development of an automated ECG interpretation and classification system would first require R&D of an effective technique to accurately and completely capture the cardiac information from the ECG signal. Such information would include the duration, amplitude and level of the component waves as well as their inter-component interval etc. Then, an intelligent decision-making system would have to be developed to analyze the recorded feature vectors and diagnose the relevant heart pathologies. The development of a reliable decision-making system would be influenced by three fundamental factors, including accurate data, pertinent KB and appropriate problem-solving scheme.

In the proposed model, Hermite based orthogonal polynomial decomposition (OPD) is explored for ECG interpretation and support vector machines (SVM) is deployed for developing ECG classification algorithm. Hermite model is adapted here for its unique benefits. Its orthogonal property allows no redundancy between feature vectors, the width parameter provides an efficient estimation of beats with large differences in QRS duration and finally the series expansion of Hermite polynomials can offer a concise representation of ECG signal. In addition, SVM is an advanced tool which has been mainly used

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

for solving classification and pattern recognition problems with high dimensional data. In contrast to ANN, the reasoning of SVM is much easier to be comprehended, a feature which is specifically useful in the design of *e*-home healthcare system. In the following section the proposed ECG interpretation and classification model using Hermite model and SVM will be described in details.

Feature Extraction using Hermite Function Decomposition

In Hermite's decomposition, the QRS complex of ECG signal $x(t)$ is expanded into a linear combination of Hermite polynomials.

$$x\left(t\right) = \sum_{n=0}^{N-1} a_n \varphi_n \left(t, \sigma\right) \tag{13}$$

$$\varphi_n \left(t, \sigma\right) = \frac{1}{\sqrt{\sigma 2^n n \sqrt{\pi}}} e^{-t^2/2\sigma^2} . H_n \left(t / \sigma\right) \tag{14}$$

Here, $H_n(t,\sigma)$ are the Hermite polynomials and the width parameter σ is proportional to the duration of QRS complex. The coefficients a_n indicates the morphological similarity between the ECG segment and various Hermite polynomials. The optimum value of coefficients a_n and width b are obtained by minimizing the error e_r between the input and synthesized signal.

$$a_{i+1}^n = a_i^n - 2 \cdot \mu_a \cdot \frac{\partial E_i}{\partial a_i^n} \tag{15}$$

$$b_{i+1} = b_i - 2 \cdot \mu_b \cdot \frac{\partial E_i}{\partial b} \tag{16}$$

$$e_r = \sum \left(x\left(t_i\right) - \sum_{n=0}^{N-1} a_n \varphi_n \left(t_i, b\right) \right)^2 \tag{17}$$

The QRS complex is centered at the R peak with a 200ms window. This size is obtained by long-term exploration by considering that both normal (60 ~ 100ms) and abnormal beats (<200ms) are covered. Validation of the selected window size was examined on arrhythmia dataset, as the length of such ECG signal would normally vary. The results show that with 10 sets of Hermite coefficients a_n and width parameter σ, a good reconstruction of the curve with most important features could be attained.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ECG Classification with Support Vector Machines

SVM belongs to a family of generalized linear classifier, which was originally designed for binary classification. However, recent studies show the theory being extended for solving multi-class classification problems (Aiolli & Sperduti, 2005; Hsu & Lin, 2002). The more commonly adapted approach views it as a collection of binary classification problem and decomposes it to a series of two-class problems. In the proposed model, one-against-all (OAA) method is adapted to execute SVM multi-class classification. In this approach, a set of binary classifiers are trained to be able to separate each individual class from other.

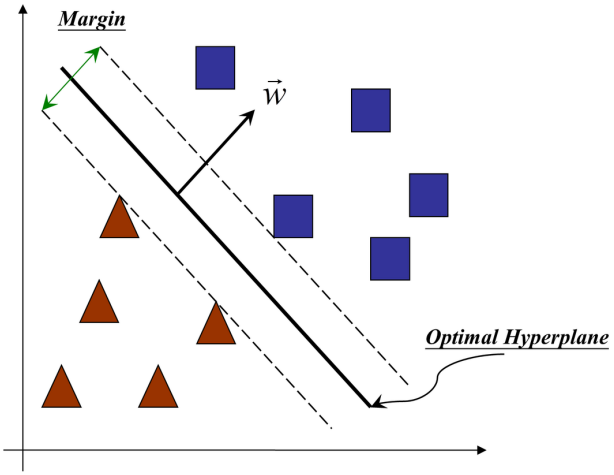
SVM classifier basically relies on the optimal hyperplane on the decision surface so that the geometric margin between two classes is maximized as shown in Figure 10. Here, the *N* set of samples be denoted as $\left(\bar{x}_i, y_i\right)$, $\bar{x}_i \in R^n$ and $y_i \in \{-1,+1\}$ are the training sample and class label respectively. The hyperplane is denoted as $f\left(\bar{x}\right) = \left\langle \bar{w}, \bar{x} \right\rangle + b = 0$, where $f\left(\bar{x}\right) > 0$ for one class and $f\left(\bar{x}\right) < 0$ for another. $\bar{w} \in R^n$ is normal vector to the hyperplane, $\left|b\right| / \left\|\bar{w}\right\|$ is the perpendicular distance from the hyperplane to the origin and $\left\|\bar{w}\right\|$ is the Euclidian norm of *w*. To obtain the hyperplane to distinguish the two classes with largest margin, the following conditions are met.

$$\left\langle \bar{w}, \bar{x} \right\rangle > +b \geq +1 \text{, } y_i = +1 \tag{18}$$

$$\left\langle \bar{w}, \bar{x} \right\rangle + b \leq -1 \text{, } y_i = -1 \tag{19}$$

As shown in Figure 12, let $d^+(d^-)$ be the shortest distance from the hyperplane to its closest positive (negative) instance. The perpendicular distance between the origin and two hyperplanes represented by Equations 18 and 19 are given by $\left\|1-b\right| - \left|-b\right\| / \left\|\bar{w}\right\|$ and $\left\|-1-b\right| - \left|-b\right\| / \left\|\bar{w}\right\|$ respectively. The margin $(d^+ + d^-)$ is equal to $2 / \left\|\bar{w}\right\|$. Smaller the value of $\left\|\bar{w}\right\|$, larger is the margin. With Equations 18 and 19, the following set of inequalities are obtained to increase the separation of the margin.

Figure 12. Support vector machines classifier



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$y_i \cdot \left(\left\langle \overline{w}, \overline{x} \right\rangle > +b \right) \geq 1, \forall i \tag{20}$$

In addition, kernel-based nonlinear SVM function could project the samples to a feature space of higher dimension via a nonlinear mapping function φ. Gaussian, polynomial and sigmoidal functions are some of the more frequently used kernel functions, which can be defined as

$$K\left(x_i, x_j\right) = \varphi\left(x_i\right) \cdot \varphi\left(x_j\right).$$

In order to maximize the distance between the margin, minimized values of $\left\|\overline{w}\right\|$ should be computed with respect to $\overline{w}$ and *b,* with the constraint of inequalities in Equation 20. Thus, with such complexity of optimization process involved, with Equation 21 the problem is converted to minimized Lagrangian function in primal space of parameters $\overline{w}$ and *b.*

$$L_P(w, b, \alpha) = \frac{1}{2}\left\|\overline{w}\right\| - \sum_{i=1}^{N} \alpha_i \left[y_i \left(< \overline{w}, \varphi\left(\overline{x}_i\right) > +b \right) - 1 \right] \tag{21}$$

where α_i is the Lagrangian multiplier and the derivative of primal Lagrangian function with respect to $\overline{w}$ and *b* as follows.

$$\frac{\partial L_P}{\partial b} = 0 \quad \Rightarrow \quad \sum_{i=1}^{N} \alpha_i \cdot y_i = 0 \tag{22}$$

With above equations, the dual Lagrangian formula $L_D(\alpha)$ can be computed as in Equation 23.

$$L_D(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K\left(\overline{x}_i, \overline{x}_j\right) \tag{23}$$

To find the optimal hyperplane, the dual Lagrangian $L_D(\alpha)$ must be maximized with respect to the non-negative multiplier ($\alpha_i \geq 0$). The dual Lagrangian can thus be expressed in terms of training data on the scalar products of input patterns ($x_i \bullet x_j$), a property that would be useful for analyzing nonlinear decision boundaries.

Through solving the dual optimization problem, the parameters $\overline{w}^{opt}$ and b^{opt} of the optimal hyperplane can be obtained as follows.

$$\overline{w}^{opt} = \sum_{i=1}^{N} \alpha_i^{opt} y_i \varphi\left(\overline{x}_i\right) \tag{24}$$

$$b^{opt} = \frac{1}{N_{SV}} \left[\sum_{s=1}^{N_{SV}} \left(\frac{1}{y_S} - < \overline{x}, \overline{w}^{opt} > \right) \right] \tag{25}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Now, with OAA approach to execute multi-class classification by separating each class from other, the radial basis kernel

$$K\left(x,y\right)=\exp\left(-\frac{\gamma\cdot\left(x-y\right)^{2}}{2\Delta^{2}}\right)$$

is chosen as the kernel function with $\gamma=2$ and $\Delta=1$. The i^{th} SVM is trained with all the training examples of the i^{th} class with positive labels and all the others with negative labels. Subsequently in an iterative process, the multi-class classification is achieved. All the values for SVM multi-class classification are determined by a series of numerical experiments on the learning and testing data to achieve the best generalization ability.

Intelligent Diagnosis System Using HS

HS by auscultation is one of the most traditional biomedical signal, as indicated by fact that the stethoscope was invented in year 1816 and has been carried and used by physicians ever since. In fact, CVD diagnosis by HS could be traced back even before the invention of stethoscope, when physicians placed their ears against patient's chest to listen to cardiac abnormality.

Even today in modern society, with more sophisticated medical techniques such as ECHO, ECG, ballistocardiogram (BCM) etc., available, HS examination through stethoscope still serves as the cheapest primary heart health screening method. Many CVD, especially the valve related diseases such as mitral valve prolapse which sometimes show no physiological symptoms are often detected during a routine examination by a general practitioner (GP) through HS examination using stethoscope. However, such diagnosis may initially go unnoticed if the GP is not experienced or confident in cardiac auscultation. An automatic analysis/diagnostic tool would thus be desirable to either implement in *e*-home healthcare system or assist GPs in diagnosing HS signals.

Exploiting the benefit of multiple technological advancements, research over the past decade has focused optimistically on the development of such tools. HS being a non-stationary time-dependent signal, researchers explored into various time-variant methods, which include the time-dependent models, time-frequency and time-scale analyses methods (Staszewski & Robertson, 2007). The time-variant approaches offered powerful solution to both de-noising (Zhao & Cao, 2009) and HS characterization (Tomas & Roncevic, 2009). In fact, these time-variant methods also extended to the development of diagnostic models. The predominant models in literature are based on HS segmentation, feature extraction and classification (Fu et al., 2010). In such approach the HS signals are first segmented into individual HS components, S1, S2, S3, S4 and murmurs. The feature vectors such as duration of individual frequency components, peak intensity, time interval etc., are extracted and then classified by relating the features to various pathological conditions. HS segmentation is done based on referencing ECG signal, time domain characteristics, frequency domain characteristics, Shannon energy etc. Time-variant methods such as S-transforms, short-time Fourier transform (STFT), wavelet transform (WT) etc., are used for feature extraction. And finally, for classification, researchers mainly adapted artificial intelligence (AI) technologies, which include hidden Morkov models, expert systems (ES), neural networks (NN) etc.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In the proposed approach, discrete wavelet transform (DWT) algorithm is applied in multi-resolution analysis (MRA) for de-noising the HS signal, envelogram for segmentation, mel-scale frequency cepstral coefficients (MFCC) algorithm is applied for HS feature extraction and dynamic time warping (DTW) is adapted for HS signal classification and CVD diagnosis. The proposed method will be further elucidated in the following sections.

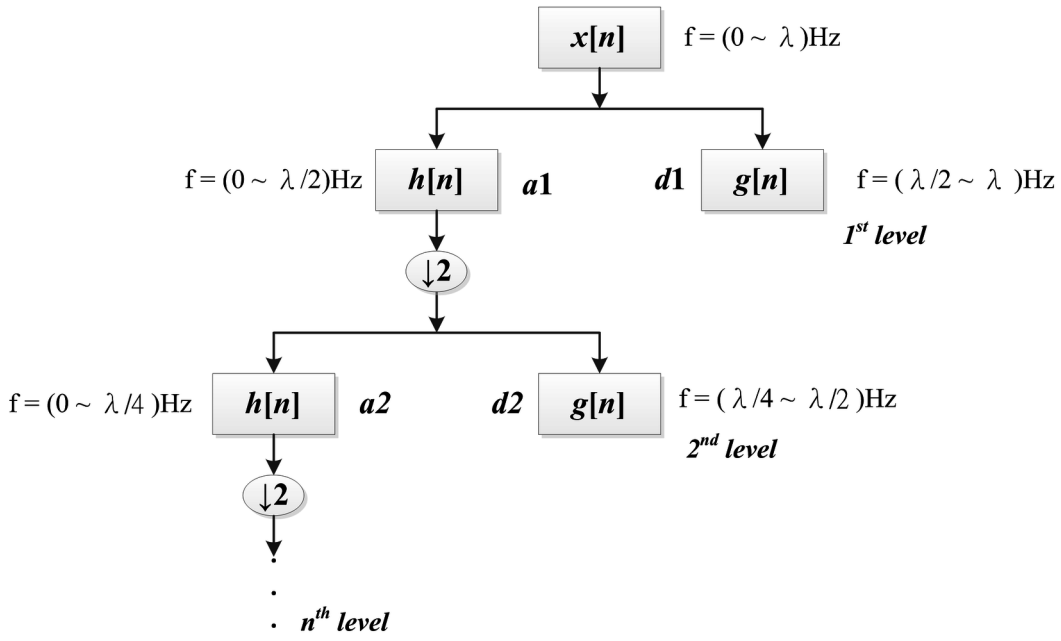
HS De-Noising with DWT in MRA

HS signal de-noising is considered to be an essential signal pre-processing step in developing computer assisted heart auscultation model. Many sources of noise including the ambient noise, lung and breath sound, fetus heart sound (if the subject is pregnant), stomach rumble, muscle tremors, noise from the recording device etc., can pollute the HS signal. As the frequencies of these signals overlap with the HS signal frequency, de-noising the HS signal without the loss of useful information has been a great challenge.

Literature shows that WT has served as a powerful tool for noise reduction (Zhao & Cao, 2009). In the proposed model, MRA algorithm based on DWT is proposed to decompose the HS signal at discrete intervals of time and scale to subsequently eliminate the noisy signal from real useful information.

A DWT is executed by using Mallat's algorithm (Mallat, 1989), which states that 'any conjugate mirror filter would characterize a wavelet that generates an orthonormal basis of $L^2(R)$'. A fast DWT is executed by cascading these conjugate mirror filters. The DWT signal decomposition can be obtained using digital FIR filters with MRA theory. The signal decomposition using DWT can be obtained as shown in Figure 13.

Figure 13. Signal decomposition using Discrete Wavelet Transform



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The signal obtained after MRA with DWT decomposition can be expressed as in Equation 26.

$$f\left(x\right)=\sum_{k}W_{\varphi}\left(j_{0},k\right)\varphi_{j_{o},k}\left(x\right)+\sum_{j=0}^{J-1}\sum_{k}W_{\psi}\left(j,k\right)\psi_{j,k}\left(x\right) \tag{26}$$

where, $\varphi(x)$ and $\psi(x)$ are scaling and wavelet function respectively; x=0,1,2,…,N is data size of original signal; *j, k* are indexes of lattice which corresponds to discrete scale step and discrete translation step respectively; constant J is the level number of the scaling coefficients. Low frequency coefficients W_{φ} and high frequency coefficients W_{ψ} can be expressed as in Equations 27 and 28.

$$W_{\varphi}\left(j_{0},k\right)=\sum_{x}f\left(x\right)\varphi_{j_{0},k}\left(x\right) \tag{27}$$

$$W_{\psi}\left(j,k\right)=\sum_{x}f\left(x\right)\psi_{j,k}\left(x\right) \tag{28}$$

With the DWT decomposition, the signal can be distributed in different frequencies band, thus the noise in specific band can be identified and eliminated using signal reconstruction with inverse DWT (IDWT). However, as some of useful information could be mixed with the noise signal in the specific frequency band, wavelet coefficient thresholding can be applied to eliminate only the noise without the loss of useful information. A brief review of the wavelet thresholding methods are presented in (Mallat, 1989). In the proposed model, the orthogonal wavelet Daubechies (db6) is chosen for wavelet decomposition and Rigorous Sure threshold estimator is used for defining the thresholding vectors.

HS Segmentation with Envelogram

HS component segmentation refers to partitioning of the HS signal into cardiac cycles and then further segmenting the individual HS components namely S1, S2, S3, S4, murmur and the systolic and diastolic intervals. HS component segmentation are basically done using two approaches:

1. With external reference to ECG and Carotid pulse (CP)),
2. Direct HS localization using signal processing techniques.

In the first approach, the relationship between the ECG, CP and HS signal is analyzed in time domain to detect the HS components and segment them subsequently. The ECG gating approach can be termed as the most robust segmentation method. However the main difficulty in this approach is that both HS and ECG sample have to be acquired from the patient simultaneously. Such complicated procedure of acquiring both ECG and HS signal, given the fact that ECG recording requires many electrodes attached to the subject would not be considered optimal in the design of *e*-home health care system.

This led researcher to explore the second approach, direct HS localization using signal processing techniques. The important step in this approach is to transform the signal to specific domain and then localize the HS components. Various transformation methods have been presented in literature which

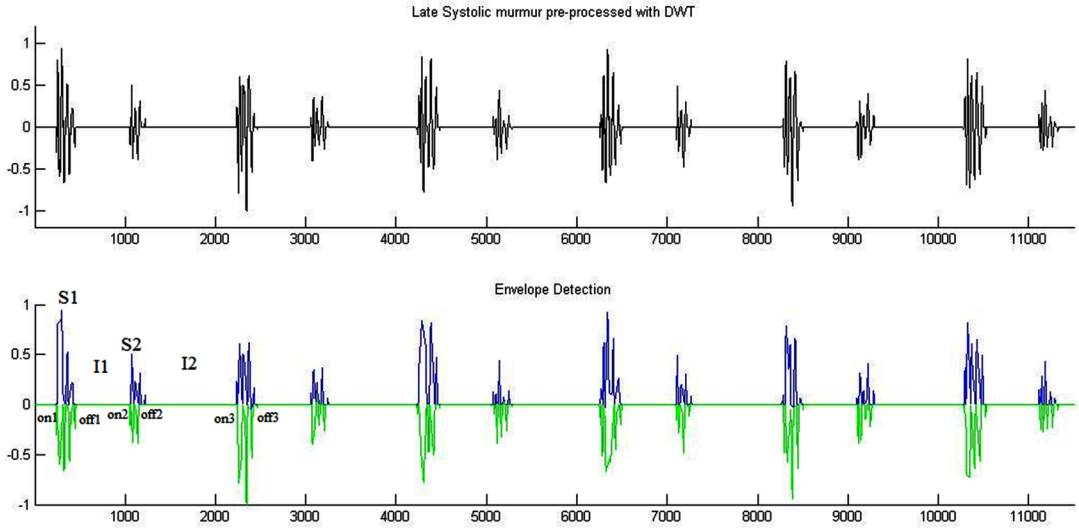
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

could be majorly classified as time-domain, frequency-domain, envelogram, matching pursuit and adaptive filtering method. Among these, envelogram approach is identified to be the most popular in direct localization of HS signal components. In this approach, envelope of HS signal is estimated by detecting and smoothly connecting the peak points of the signal. Then through analysis of the obtained envelope, the cardiac components of HS signal are identified and subsequently segmented for further analysis. Few examples which use envelogram for HS segmentation are the Shannon energy (Chauhan et al., 2008), homomorpic filtering (Gupta et al., 2007), etc.

The main point to be noted while estimating the envelope of the signal is that the algorithm has to attain a balance between loosely connecting only the prominent peak points to obtain a rough envelope of the signal or closely connecting all the detected peaks to result into an over fitted envelope that almost resembles the original signal itself. Thus, the envelogram estimation algorithm should identify the essential peaks of the signal and smoothly connect them to obtain the envelope of the signal for further segmentation.

Thus, in the proposed approach the piecewise cubic spline interpolation theory (Marsden, 1974) is investigated to attain a smooth curve of the amplitude envelope that approximately matches the main peaks of HS waveform. With this theory, a series of unique cubic polynomials are fitted between each extrema position of the HS waveform with the stipulation that the estimated envelope is continuous and smooth. This method provides a reliable estimation of the envelope that closely follows the sudden variations in amplitude of waveform and avoids ripples in more stable regions. The envelope detection of pathological HS signal with late systolic murmur using piecewise interpolation function is shown in Figure 14. The HS components are detected by measuring the interval difference between the boundary of the slope where the upper and lower envelope intersect. Based on the fact that the systole (I1) is much shorter than diastole (I2), S1 and S2 components of HS signal are easily identified and subsequently segmented.

Figure 14. Envelope detection of HS signal using piecewise interpolation function



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

HS Feature Extraction using MFCC

Generally for feature extraction, signal processing is carried out to convert the raw signal to some type of parametric representation. The parametric representation is then processed to represent the features of the signal which are then used for further recognition or classification purpose. MFCC is a parametric representation which to a certain degree simulates the properties of the human auditory system. Thus, the derived MFCC coefficients could attain more sound characteristics from low frequencies and fewer from high frequencies. MFCC was mainly developed for analyzing auditory signals and has been well applied for automatic speech recognition (Davis & Mermelstein, 1980). Also, recent studies show that MFCC has been frequently applied for HS analysis and has attained high classification accuracy (Gupta et al., 2007).

Figure 15 shows the functional block diagram for computing MFCC. As described above, HS signal is first de-noised, segmented and normalized before treated for computing MFCC. Now, once the HS signal is pre-processed, the segmented HS component is processed using framing. In this step the HS signal is divided into frames of 8~15ms length, with each overlapping with a shift of 1/2 of the frame. Then hamming window is applied to frames to minimize the spectral leakage and obtain a smooth transition between frames (Fu & Wang, 2010).

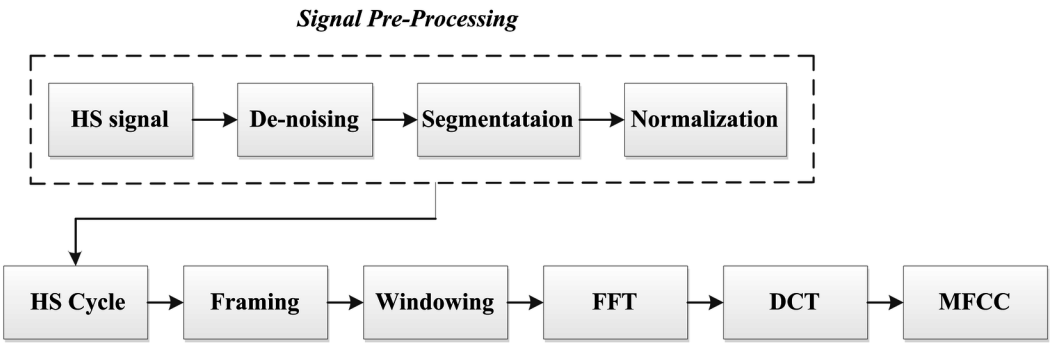
FFT is then computed on each window samples to extract the frequency components of the signal from the time domain. Subsequently, MFCC coefficients are obtained from a Discrete Cousine Transform (DCT) of the logarithm of the short-term energy spectrum expressed on a mel-frequency scale. The desired mel-frequency scale is obtained using triangular band pass filters. Equation 29 expresses the relationship between mel-frequency and actual frequency:

$$Mel(f) = 1125 \ln\left(1 + \frac{f}{700}\right) \tag{29}$$

Here, *f* is the physical frequency in Hz and *Mel(f)* is the perceived frequency in mels.

A DCT is applied to the logarithm of the filter bank outputs to calculate the MFCC as described in Equation 30.

Figure 15. Functional block diagram for computing MFCC



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$C_n = \sum_{k=1}^{N} E_k \cos\left[n\left(k-\frac{1}{2}\right)\frac{\pi}{K}\right] \tag{30}$$

Here, C_n denotes the n-th MFCC coefficient, E_k is the logarithm output of the k-th filter bank channel, K is the number of filter bank channels, n=1,2, …K and N is the number of filter banks.

HS Classification using Dynamic Time Warping

DTW algorithm aims at comparing the similarity between two sequences of data points which could vary in time or speed. DTW was first introduced in 1960s (Bellman & Kalaba, 1959) and then extensively explored for speech recognition in 1970s (Sakoe & Chiba, 1978). Subsequently, it has been employed in countless applications including online biometric recognition (Faundez-Zanuy, 2007), computer animation (Müller, 2007), protein sequence analysis (Vial et al., 2009) etc.

DTW works in time domain by iteratively 'warping' between the two sequences axis until an optimal match is found. Hence, the main advantage of DTW over other methods such as Euclidean distance is that the sequence of data under analysis need not be of equal length. It allows elastic shifting of time axis to accommodate sequences which are similar but out of phase.

As an example, for two sequences A=(a_1,a_2,…,a_m,…,a_M) and B=(b_1,b_2,…,b_n,…,b_N) of varying lengths M and N respectively, a_m is the m^{th} element of sequence A and b_n is the n^{th} element of sequence B. The distance between the two points can be represented as in Equation 31, which tries to find the shortest path between the two points.

$$d_{m,n} = d(a_m, b_n) = (a_m - b_n)^2 \tag{31}$$

As shown in Figure 16, the warping path W would measure the continuous set of matrix to map the distance between the two sequences A and B.

Now, as brute force approach of finding the shortest warping path would not be efficient, DTW algorithm uses two specific features that include some constraints and dynamic programming.

First to limit the number of paths to explore in mapping the distance between two sequences, DTW follows some reasonable constrains, which include monotonicity, continuity, boundary, warping window and shape. These constraints are elaborately described in (Kurbalija et al., 2011).

With the k^{th} element of warping path W is defined as w_k=$(m,n)_k$.

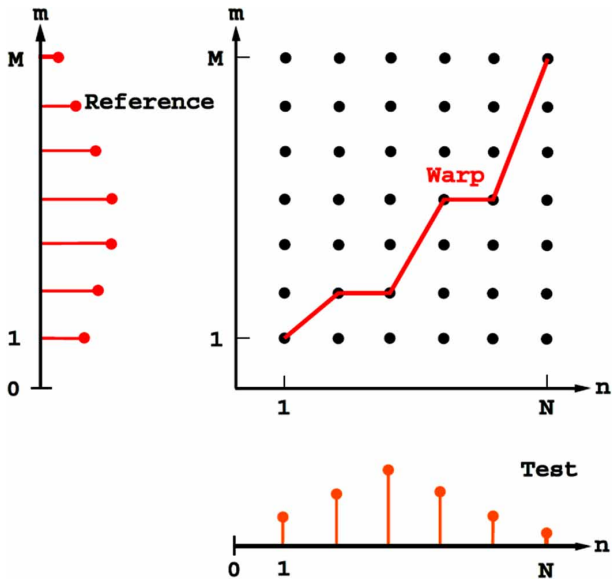
$$\begin{aligned} &W = w_1, w_2, \cdots, w_k, \cdots, w_K \\ &\max(M,N) \le K \le M+N-1 \end{aligned} \tag{32}$$

The total cost of the warping path between A and B can now be defined as in Equation 33.

$$C_W\left(A,B\right) := \sum_{k=1}^{K} w_k \tag{33}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. Measure of warping path from A to B using DWT



Secondly, DTW uses dynamic programing using Equation 34 to find the best alignment in a recursive manner.

$$c(m,n) = d(a_m, b_n) + \min\left\{c(m-1,n-1), c(m-1,n), c(m,n-1)\right\} \tag{34}$$

The optimal warping path between *A* and *B*, which holds minimal total cost of all possible warping paths can be defined as in Equation 35.

$$DTW\left(A,B\right) := \min\left\{C_W\left(A,B\right)\right\} = min\sum_{k=1}^{K} w_k \tag{35}$$

Testing Results of Midstream Module

As the IMDS in the midstream utilizes three non-invasive biological signals including SPG, ECG and HS, each of the proposed and designed intelligent diagnosis models will be tested individually.

Three indices, namely the sensitivity (SE), specificity (SP) and accuracy (A) defined in Equations 36, 37, and 38 are used to evaluate the performance level of the proposed IMDS.

$$SE = \frac{True_Positive}{True_Positive + False_Negative} \times 100\% \tag{36}$$

$$SP = \frac{True_Negative}{True_Negative + False_Positive} \times 100\% \tag{37}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$A = \frac{True_Positive + True_Negative}{True + False} \times 100\% \tag{38}$$

First to test the intelligent diagnosis model using FHNNs designed for SPG analysis, records from the site-sampled data sets obtained from two hospitals in China are tested. To carry out the validation process, 52 CHD, 48 HT, 63 HL and 50 normal records are selected to test the performance level of the trained FHNNs.

The validation results of proposed model shown in Figure 17 indicate that 'P-P-L' FHNNs performs better than all other schemes of fusing FNNs (denoted as F), FPNNs (denoted as P) and FLVQ (denoted as L).This can be evidenced for classifying all the three types of CVD and the mixed CVD taken for testing.

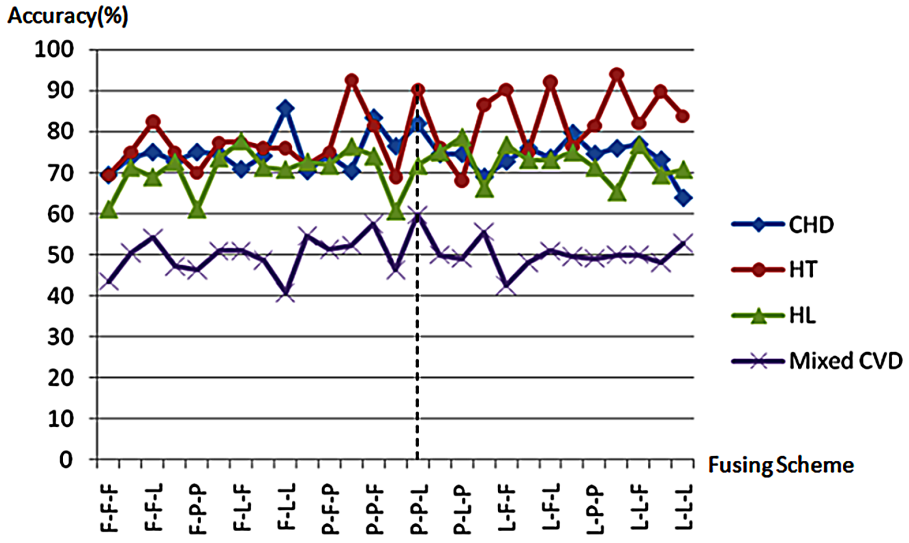
Secondly, to validate the proposed ECG classification using Hermite based orthogonal polynomial decomposition and SVMs, ECG records of normal sinus rhythm and seven types of arrhythmia are obtained from the MIT-BIH Arrhythmia database (Mark & Moody, 1998). The validation is made based on both intrapersonal and interpersonal scheme. Here intrapersonal (interpersonal) testing represents the interactive influences contained within (between) the person. For testing the intrapersonal and interpersonal scheme, a total of 6138 and 3174 beats are segmented from the acquired ECG records.

Tables 4 and 5 show the testing results of intrapersonal and interpersonal classification.

The testing results from Table 4 and 5 show that the proposed method is competitive in classifying ECG samples based on intrapersonal scheme, however, further improvement in the algorithms can be made in classifying based on the more complicated interpersonal scheme.

Finally, to validate the proposed HS classification method using MFCC and DTW, HS samples with varying pathologies are acquired from eGeneralMedical.com, (n.d), USA; University of Michigan heart sound and murmur library, (n.d); University of Washington School of Medicine, (n.d); University of Dundee, (n.d) etc., databases.

Figure 17. Comparison of diagnostic accuracy between diverse fusing schemes



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The validation result of intrapersonal classification

ECG Beat Type	Sensitivity (%)	Specificity (%)
Normal sinus rhythm	99.42	99.92
Ventricular premature beat	95.34	99.49
Left bundle branch block	93.80	99.73
Right bundle branch block	98.76	99.59
Ventricular flutter wave	92.50	99.78
Fusion of paced and normal beat	93.50	99.97
Fusion of ventricular and normal beat	95.24	99.55
Atrial premature beat	97.52	99.52
Total	**95.76**	**99.70**

Table 5. The validation result of interpersonal classification

ECG Beat Type	Sensitivity (%)	Specificity (%)
Normal sinus rhythm	80.61	99.42
Ventricular premature beat	79.95	99.68
Left bundle branch block	62.40	92.93
Right bundle branch block	85.60	97.23
Ventricular flutter wave	N/A*	99.81
Fusion of paced and normal beat	N/A	97.35
Fusion of ventricular and normal beat	N/A	94.90
Atrial premature beat	N/A	96.06
Total	**77.14**	**97.17**

*N/A denotes the case of zero denominators.

Table 6 shows the performance evaluation of the proposed method of HS analysis and classification.

The testing results in Table 6 show that the proposed method of HS analysis and classification is competitive and can be reliably adapted in the design of *e*-home healthcare system.

UPSTREAM FUNCTIONAL MODULE

With the proposed architecture of distributed *e*-home healthcare system, so far it can be recognized that already the down- and mid-steam could independently offer reliable monitoring of the cardiac status of the patient at the comfort of their home environment. Yet, to design a more reliable *e*-home healthcare system as a whole, it is important to have element of intervention by the hospital during emergency situations. Thus, as depicted in Figure 1, the upstream module is introduced. The upstream basically consists of two sub-systems, namely customized information management system (CIM) and top-level intelligent CVD monitoring and diagnosis system (TICMD) (Tang, 2014). While CIM works with UUS

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. The performance evaluation of HS analysis and classification

HS Pathology Type	Sensitivity (%)	Specificity (%)	Accuracy (%)
Normal Sound	97.7	92.3	93.2
Aortic Insufficiency	100	100	100
Aortic Regurgitation	79.3	99.3	98.2
Aortic Stenosis	79.5	99.0	95.8
Atrial Septal Defect	82.3	100	99.4
Mitral Regurgitation	77.0	99.5	96.8
Mitral Stenosis	83.7	97.6	96.2
Mitral Prolapse	93.1	99.8	99.4
Pulmonary Stenosis	87.5	93.6	93.2
Pericarditis	92.9	100	99.8
Tricuspid Regurgitation	91.2	100	99.4
Ventricular Septal Defect	100	100	100
Gallop	50.9	99.3	94.2
Total	**82.9**	**98.6**	**97.4**

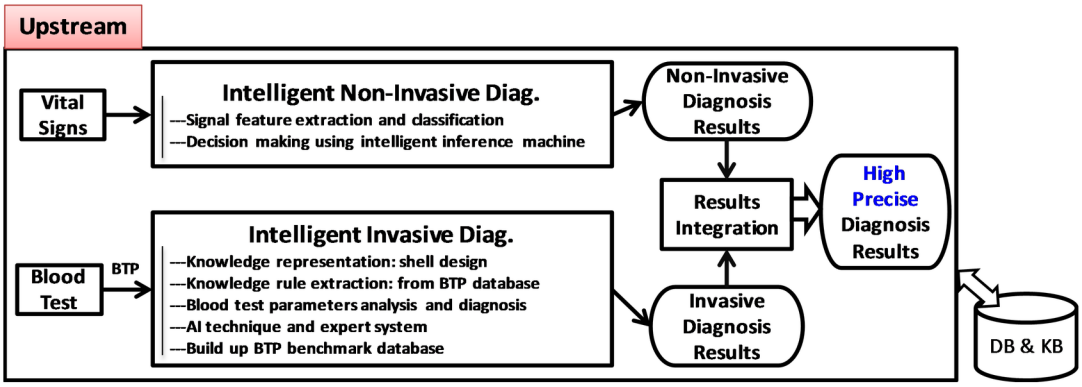
to manage the help requests from lower streams, TICMD performs high precision diagnoses with detailed explanations and suggestions. As any failure of diagnosis in lower streams due to resource limitation would seek for support from a stronger KB and diagnosis model in upstream, a special attention has to be paid in the design of TICMD. The designed TICMD should provide absolutely authoritative and eminently trustworthy reference for the overall *e*-home healthcare system.

Thus, the proposed functional architecture of TICMD for upstream as illustrated in Figure 18, employs both non-invasive multi-vital signs including SPG, ECG, HS and invasive blood test parameters (BTPs) extracted from routine blood test to diagnose various CVD.

Note that different from biomedical signals interpretation, the intelligent diagnosis and computer-aided analysis of BTPs are seldom exploited in literature. Thus, to construct such a model, 241 site-sampled BTPs records from 241 patients with different CVD are acquired from Shaoshan People's Hospital, China. Based on these records, a dedicated DB is constructed. In order to provide standard for BTPs related studies and guarantee clarity, consistency, extensibility, and interoperability of DB, the frame structure and shell design is employed to develop the DB. Also, a KB system is devised to explain the mapping relationship between certain CVD and related BTPs. Through cooperation with medical specialists and clinical validations, 26 KB rules with specialized architecture are pre-defined. Grounded on the KB rules and input BTPs in DB, an intelligent inference engine is developed to conclude the accurate diagnostic results with detailed explanations with corresponding warning messages. After obtaining both invasive and non-invasive diagnostic results, final decisions can be integrated by considering the influence factor of different parameters for different diseases.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. Functional architecture of upstream



Testing Results of Upstream Module

To validate the designed TICMD, 200 site-samples, BTPs records containing 26 different types of CVD obtained from Shaoshan People's hospital were tested. The disease distribution of all the acquired records are listed in (Tang, 2014).

The quantitative judgment and evaluation on diagnosis performance of proposed TICMD is evaluated by computing the diagnostic accuracy (*DA*) using Equation 39.

$$DA\left(\%\right) = \frac{N_{sim}}{N_{tot}} \times 100\% \tag{39}$$

where, N_{sim} is the number of similar diagnostic results obtained against the actual clinical results from medical professionals and N_{tot} denotes the total number of CVD types present in the acquired database.

With the proposed TICMD, the testing samples are diagnosed automatically. The recorded diagnostic results are compared with doctor's clinical results to obtain the performance index *DA*. The distribution of *DA* for all tested records is plotted in Figure 19.

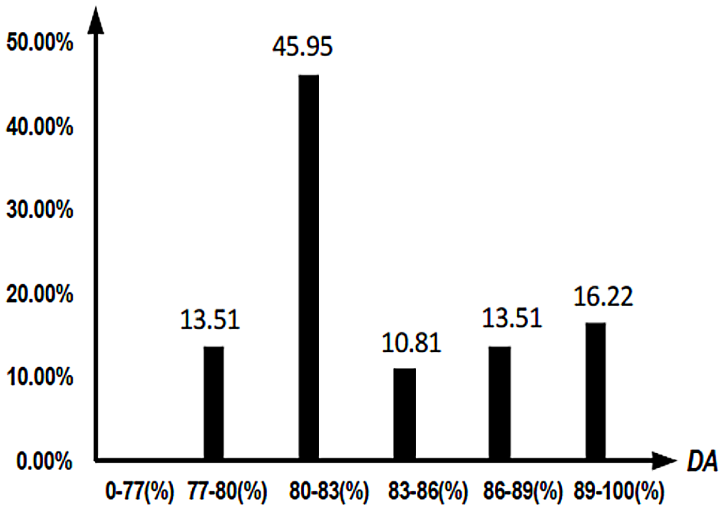
It is observed form Figure 19 that for most of the records, the *DA* is greater than 80% and the averaged *DA* obtained is 84.38%. Such results show that the proposed TICMD is reliable and the achieved performance is adequate for the implementation of *e*-home healthcare system.

UPLINK, UPDATE, AND SYNCHRONIZATION SYSTEM

As a result of intrinsic resource limitation, DB and KB in mid- and down-streams are rather restricted to handle the numerous CVD and over-complicated medical diagnosis. Therefore, UUS system is specialized as backyard on-line assistance to support the functionality of versatile distributed CVD diagnosis and enable remote communication between thousands of home users, healthcare center, and hospital server.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 19. Distribution of DA for the testing samples



As depicted in Figure 20, the architecture of UUS contains five parts, namely help request delivery, help request pretreatment, customized help request management, disease driven rule base update and synchronization. There among, help request delivery and rule base updated are performed in local devices in mid- and down-streams while other processes are conducted in server of upstream.

Whenever the users need uplink help, they can send a customized help request to server. During local self-diagnosis, the diagnostic results, detailed explanations and suggestions are displayed on GUI. Meanwhile, a help enquiry button is displayed to see whether the user wants to confirm the results. Once the user chooses to ask for help, the type of help will be obtained from the local results which could be normal or critical. After that, the acquired biomedical signal, local diagnostic results, personal information, as well as the encoded help type could be combined to form the help information. After compressing the help information, a customized message file is obtained and transmitted to the selected server.

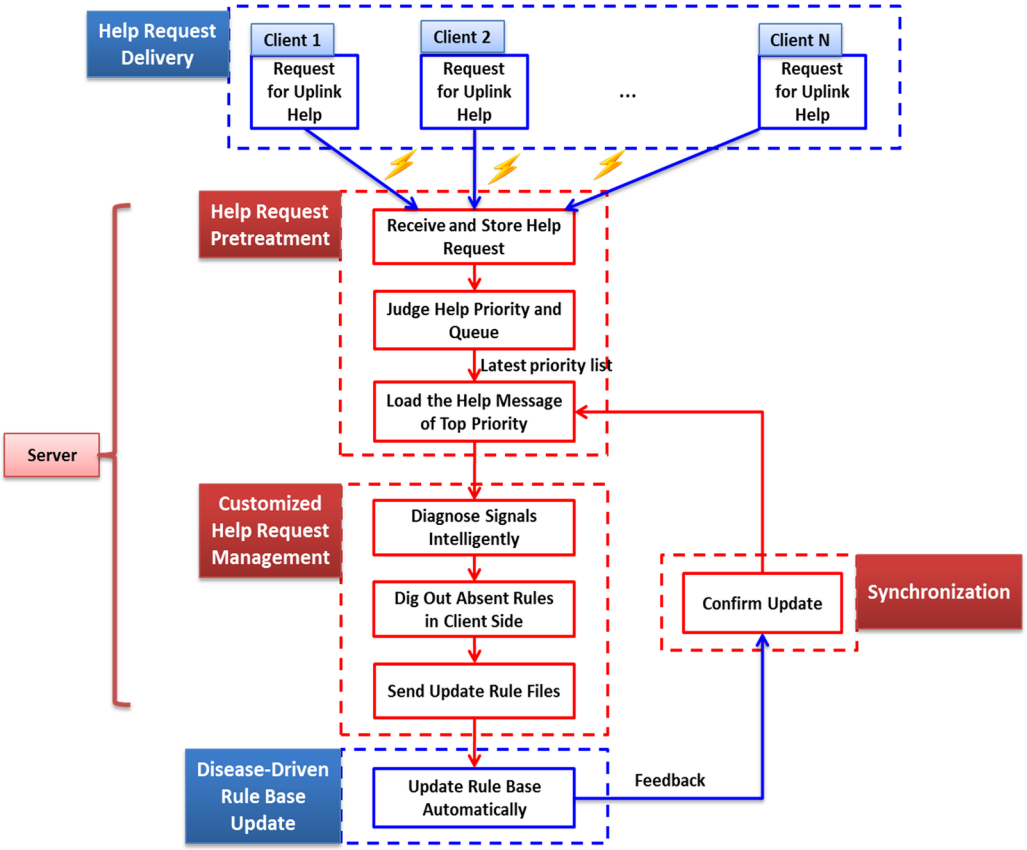
On receiving the help message, the server would first store and decompress it to get the original information from which the help type and user ID are obtained. According to the level of help type, the priority list is updated where critical help is ranked higher than normal ones. If the priority of new incoming help request is higher than all the existing help requests, an interruption scheme can be activated otherwise the new request would be inserted to the waiting list. The help message with top priority is always loaded for further processing.

After pretreatment, a customized management scheme is performed. First, based on the acquired signals, top-level intelligent diagnosis in server is conducted to obtain high precise diagnostic results and the useful rules during diagnosis are recorded with four digit number. Afterwards, these diagnostic results are compared with local results. If the same results are obtained, which indicates that the local devices can handle such diseases, a confirmation feedback message is sent to local device. Otherwise, update procedures are triggered automatically. According to the client's rule numbers, server can dig out the absent rules in client's rule base. These absent rules as well as the top-level diagnostic results are then packaged and compressed to be transmitted to local devices.

In order to reduce the size of update package and conduct update process with least resource and band-width, a delta update scheme is proposed (Choi et al., 2014). Instead of sending back all the re-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 20. Functional architecture of uplink, update and synchronization scheme



lated useful rules, only the absent or outdated rules for certain disease are dug out and transmitted to local devices. Such a delta update scheme could minimize the necessary dataflow between up and lower streams in band-restricted network, and it also contributes to easy execution of local update without the need of re-installation.

Once the up-down communication link is established, the whole update process is carried on automatically and intelligently without manual intervention. Upon receiving the feedback rule update package, it is firstly decompressed to obtain the original update files, including top-level diagnostic results and absent rules. Then the local results are updated and displayed. At the same time, the update process of rule base in client devices is conducted automatically in the background by inserting absent rules into the original local rule base. Note that rule base update is the key part of UUS to customize the local KB for tailor-made diagnosis. During the update process, several unexpected situations may occur, such as sudden power off, system failure, system exit etc. Such update failure may damage the original KB file and even the overall system. Therefore, in virtue of reliable and efficient update, a master-slave KB update scheme is proposed in the design of UUS. The original local KB file is regarded as master KB, which takes the responsibility of diagnosis until the successful completion of update process. Instead of directly overwriting the master KB, a blank file is created as slave KB. By analyzing the received update package and original master KB file, up-to-date rules can be recorded in the slave KB. During the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

update process, if any interruption happens, the incomplete slave KB file is deleted without influencing the following diagnosis. After successful update, the slave KB is adopted as master KB. In addition, the original master KB is archived for backup, in case the user wants to roll back to the original version. Such dedicated update scheme guarantees the robustness and reliability of the entire system.

As the last procedure of communication, as soon as the confirmation feedback message from client is received, the corresponding synchronization procedure is triggered. Synchronization step is executed to confirm the success of update and guarantees the data consistency between server and local client devices.

FUTURE RESEARCH DIRECTIONS

In order to efficiently prevent from chronic diseases, alleviate the burden of medical personnel and decrease the cost of accessible healthcare service in the context of current medical service industry, a novel paradigm '*e*-health' plays a critical role. Ubiquitously, *e*-health is enjoying a booming development push from various scientific and technological fields, including ICT, mobile technology, WBSN, intelligent technology etc. With the aid of WBSN and embedded link devices, *e*-health provides platform for monitoring individual's health related parameters as well as processing and providing relevant information to the healthcare consumers and/or medical professionals on regular basis. On the other hand, by delivering healthcare for anyone, anytime and anywhere, *e*-health removes the location, time and other restrains and increases both the coverage and quality of healthcare services.

Undoubtedly *e*-health is accelerating towards the remarkable shift from conventional hospital-centered provision paradigm to customized individual-centered proactive approach. Now to facilitate such a change, as the future research directions, one could expect to have advanced developments in the field of *e*-health. To name a few, this could include the advances in *e*-health tools, such as personalized electronic health records, lifetime health records, medical information management system etc.; introduction of *e*-health awareness/education, such as public health education, personalized healthcare professional continuing education etc.; advances in web technology in *e*-health for remote health monitoring, including telehealth and tele-care, virtual healthcare teams and wards etc. In short, the goal of attaining the customized individual-centered proactive healthcare approach is already defined. Yet, to reach such a goal, one could expect to have research and development of various key technologies in the field of *e*-health.

CONCLUSION

This chapter introduces systematic framework architecture of a versatile wired and wireless distributed *e*-home healthcare system to enable comprehensive and timely healthcare services for patients suffering from CVD. The proposed design merges intelligent multi-vital sign acquisition, customized information management and hierarchical diagnosis in wired and wireless platform to form as a complete healthcare system that enables cardiac health self-monitoring and reserves the need of hospital services for only critical and emergency situations. Thus, the model principally addresses the key challenges of network dependence and medical personnel necessity in the existing *e*-health application and promotes patient-centric healthcare. As a concluding remark, such a system in place would not only cater the aging population in monitoring their health status but could also encourage people to adopt a preventive healthcare lifestyle and reduce the mortality rate due to cardiac failure. Moreover, though the proposed wired and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

wireless distributed *e*-home healthcare system is mainly designed for CVD monitoring and diagnosis, the methodology can be generalized and adapted for many other *e*-health services for aging population, including monitoring and diagnosis of diabetes, Parkinson's disease, dementia etc.

ACKNOWLEDGMENT

This research was supported by the Science and Technology Development Fund (FDCT) of Macau S.A.R with project ref. No. 016/2012/A1 and also by Research Committee of University of Macau under Grant No. MYRG2014-00060-FST.

REFERENCES

Adibi, S. (2012). Link technologies and blackberry mobile health (mhealth) solutions: A review. *IEEE Transactions on Information Technology in Biomedicine*, *16*(4), 586–597. doi:10.1109/TITB.2012.2191295 PMID:22453643

Aiolli, F., & Sperduti, A. (2005). Multi-prototype support vector machines. *Journal of Machine Learning Research*, 817–850.

Asgary, R., & Mohammadi, K. (2005). Using fuzzy probabilistic neural network for fault detection in MEMS. *Proceedings of the 5th International Conference on Intelligent Systems Design and Applications (ISDA)*.Wroclaw, Poland:IEEE. 10.1109/ISDA.2005.96

Bagher-Ebadian, H., Soltanian-Zadeh, H., Setayeshi, S., & Smith, S. T. (2004). Neural network and fuzzy clustering approach for automatic diagnosis of coronary artery disease in nuclear medicine. *IEEE Transactions on Nuclear Science*, *51*(1), 184–192. doi:10.1109/TNS.2003.823047

Baghshah, M. S., Shouraki, S. B., & Kasaei, S. (2006). A novel fuzzy classifier using fuzzy LVQ to recognize online Persian handwriting. *Proceedings of the 2nd Information and Communication Technologies (ICTTA)*.Damascus, Syria: IEEE. 10.1109/ICTTA.2006.1684675

Bellman, R., & Kalaba, R. (1959). On adaptive control processes. *I.R.E. Transactions on Automatic Control*, *4*(2), 1–9. doi:10.1109/TAC.1959.1104847 PMID:16590506

Calado, J. M. F., & Sa da Costa, J. M. G. (1998). A hierarchical fuzzy neural network approach for multiple fault diagnosis. In IEE conference publication, 1498-1503. doi:10.1049/cp:19980451

Caldeira, J. M., Rodrigues, J. J., & Lorenz, P. (2012). Toward ubiquitous mobility solutions for body sensor networks on healthcare. *IEEE Communications Magazine*, *50*(5), 108–115. doi:10.1109/MCOM.2012.6194390

Chauhan, S., Wang, P., Lim, C. S., & Anantharaman, V. (2008). A computer-aided MFCC-based HMM system for automatic auscultation. *Computers in Biology and Medicine*, *38*(2), 221–233. doi:10.1016/j.compbiomed.2007.10.006 PMID:18045582

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chiang, H. S., Binshan, L., & Lin, S. B. (2010). An embedded mobile ECG reasoning system for elderly patients. *IEEE Transactions on Information Technology in Biomedicine*, *14*(3), 854–865. doi:10.1109/TITB.2009.2021065 PMID:19403370

Choi, U. H., Ma, J. L., Guo, R., & Dong, M. C. (2014). Intelligent knowledge management system for distributed *e*-home healthcare. *International Journal of Bioscience*. *Biochemistry and Bioinformatics.*, *4*(4), 270–274.

Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *28*(4), 357–366. doi:10.1109/TASSP.1980.1163420

Heart Sound Samples. (n. d). *eGeneralMedical.com*. Retrieved from http://www.egeneralmedical.com/listohearmur.html

Faundez-Zanuy, M. (2007). On-line signature recognition based on VQ-DTW. *Pattern Recognition*, *40*(3), 981–992. doi:10.1016/j.patcog.2006.06.007

Fu, W., Yang, X., & Wang, Y. (2010). Heart sound diagnosis based on DTW and MFCC. Proceedings of the *3rd International Congress on Image and Signal Processing (CISP)*.Yantai, China:IEEE. 10.1109/CISP.2010.5646678

Gan, Z., Fu, B. B., & Dong, M. C. (2014). Interactive and wearable MIP recognition technique combine pattern recognition and spatial tracking based on SPG sampling. *WSEAS Transactions on Biology and Biomedicine*, *11*, 82–87.

Gelman, A. (2005). Analysis of variance? Why it is more important than ever. *Annals of Statistics*, *33*(1), 1–53. doi:10.1214/009053604000001048

Georgiou, V. L., & Vrahatis, M. N. (2008). Fuzzy evolutionary probabilistic neural networks. In Artificial Neural Networks in Pattern Recognition, 113-124. doi:10.1007/978-3-540-69939-2_11

Ghaffari, A., Homaeinezhad, M. R., Akraminia, M., Atarod, M., & Davaeeha, M. (2009). A robust wavelet-based multi-lead electrocardiogram delineation algorithm. *Medical Engineering & Physics*, *31*(10), 1219–1227. doi:10.1016/j.medengphy.2009.07.017 PMID:19692287

Gupta, C. N., Palaniappan, R., Swaminathan, S., & Krishnan, S. M. (2007). Neural network classification of homomorphic segmented heart sound. *Applied Soft Computing*, *7*(1), 286–297. doi:10.1016/j.asoc.2005.06.006

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, *13*(2), 415–425. doi:10.1109/72.991427 PMID:18244442

Husser, D., Bollmann, A., & Sornmo, L. (2009). Waveform characterization of atrial fibrillation using phase information. *IEEE Transactions on Bio-Medical Engineering*, *56*(4), 1081–1089. doi:10.1109/TBME.2008.2006624 PMID:19272897

Hwang, J. M., Park, J. M., & Park, S. O. (2004). Wrist-located pulse detection using IR reflection due to pulse added volume of arterial blood. *Proceedings of 3rd International Conference on Computational Electromagnetics and Its Applications (ICCEA)*.United States: IEEE.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Im, J. J., & Lessard, C. S. (1995). A study for the development of a noninvasive continuous blood pressure measuring system by analyzing radial artery pulse from a wrist. Proceedings of *IEEE 17th Annual Conference* on *Engineering in Medicine and Biology Society (EMBC),* Quebec, Canada: IEEE. 10.1109/IEMBS.1995.579454

Jatmiko, W., Rochmatullah, R., Kusumoputro, B., Sekiyama, K., & Fukuda, T. (2009). Fuzzy learning vector quantization based on particle swarm optimization for artificial odor discrimination system. *WSEAS Transactions on Systems*, *8*(12), 1239–1252.

Kara, A. (2001). Protecting privacy in remote-patient monitoring. *Computer*, *34*(5), 24–27. doi:10.1109/2.920607

Kovalerchuk, B., Triantaphyllou, E., Ruiz, J. F., & Clayton, J. (1997). Fuzzy logic in computer aided breast cancer diagnosis: Analysis of Iobulation. *Artificial Intelligence in Medicine*, *11*(1), 75–85. doi:10.1016/S0933-3657(97)00021-3 PMID:9267592

Kurbalija, V., Radovanović, M., Geler, Z., & Ivanović, M. (2011). The influence of global constraints on DTW and LCS similarity measures for time-series databases. Proceedings of the *Third International Conference on Software, Services and Semantic Technologies.* Bourgas, Bulgaria: Springer Berlin Heidelberg. 10.1007/978-3-642-23163-6_10

Kyriacou, E., Pavlopoulos, S., Berler, A., Neophytou, M., Bourka, A., Georgoulas, A., & Koutsouris, D. (2003). Multi-purpose healthcare telemedicine systems with mobile communication link support. *Biomedical Engineering Online*, *2*(7), 1–12. PMID:12694629

Lee, S. J., Kim, J., & Lee, J. H. (2011). A real-time ECG data compression and transmission algorithm for *e*-health devices. *IEEE Transactions on Bio-Medical Engineering*, *58*(9), 2448–2455. doi:10.1109/TBME.2011.2156794 PMID:21606020

Lei, W.K., Li, B.N., Fu, B.B., & Dong, M.C. (2007). An application of morphological feature extraction and support vector machines inference in automatic ECG interpreter. Proceedings of the *Sixth Mexican International Conference on Artificial Intelligence-Special Session (MICAI).*Aguascalientes, Mexico: IEEE. 10.1109/MICAI.2007.32

Lei, W.K., Li, B.N., Dong, M.C., & Vai, M.I. (2007). AFC-ECG: an adaptive fuzzy ECG classifier. *Soft Computing in Industrial Applications,* 189-199.

Li, B. N., Dong, M. C., Vai, M. I., & Un, M. P. (2005). A novel intelligent sphygmogram analyzer for health monitoring of cardiovascular system. *Expert Systems with Applications*, *28*(4), 693–700. doi:10.1016/j.eswa.2004.12.026

Liu, M., Zhang, Q., Lu, M., Kwon, C. S., & Quan, H. (2007). Rural and Urban Disparity in Health Services Utilization in China. *Medical Care*, *45*(8), 767–774. doi:10.1097/MLR.0b013e3180618b9a PMID:17667311

Ma, J. L., Cui, Y. L., & Dong, M. C. (2013). An effective low-complexity multi-vital-signs compression technique for embedded-link *e*-home healthcare. Proceedings of the *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1177-1181.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ma, J. L., Zhang, T. T., & Dong, M. C. (2015). A novel ECG data compression method using adaptive Fourier decomposition with security guarantee in *e*-health applications. *IEEE Journal of Biomedical and Health Informatics*, 986–994. PMID:25222961

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7), 674–693. doi:10.1109/34.192463

Manikandan, M. S., & Dandapat, S. (2007). Wavelet-based ECG and PCG signals compression technique for mobile telemedicine. Proceedings of *IEEE International Conference on Advanced Computing and Communications*, 164-169. 10.1109/ADCOM.2007.29

Mark, R., & Moody, G. (1998). *MIT-BIH arrhythmia database directory*. Cambridge: Massachusetts Institute of Technology.

Marsden, M. (1974). Cubic spline interpolation of continuous functions. *Journal of Approximation Theory*, *10*(2), 103–111. doi:10.1016/0021-9045(74)90109-9

Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine*, *3*(11), 2011–2029. doi:10.1371/journal.pmed.0030442 PMID:17132052

Miao, J. (1992). A computer aided method for design of noninvasive indirect measurement of arterial blood pressure. *Proceedings of Southeastcon.* Birmingham, Alabama: IEEE. 10.1109/SECON.1992.202256

Mishra, S., Bhende, C. N., & Panigrahi, K. B. (2008). Detection and classification of power quality disturbances using S-transform and probabilistic neural network. *IEEE Transactions on Power Delivery*, *23*(1), 280–287. doi:10.1109/TPWRD.2007.911125

Müller, M. (2007).DTW-based motion comparison and retrieval. *Information Retrieval for Music and Motion*, 211-226.

Niyato, D., Hossain, E., & Camorlinga, S. (2009). Remote patient monitoring service using heterogeneous wireless access networks: Architecture and optimization. *IEEE Journal on Selected Areas in Communications*, *27*(4), 412–423. doi:10.1109/JSAC.2009.090506

Oster, J., Pietquin, O., Kraemer, M., & Felblinger, J. (2010). Nonlinear Bayesian filtering for denoising of electrocardiograms acquired in a magnetic resonance environment. *IEEE Transactions onBiomedical Engineering, 57*(7), 1628-1638.

Patra, D., Ray, S., Mukhopadhyay, J., Majumdar, B., & Majumdar, A. K. (2009). Achieving *e*-health care in a distributed EHR system. *Proceedings of IEEE 11th International Conference on e-Health Networking, Applications and Services*, Sydney, NSW. IEEE. 10.1109/HEALTH.2009.5406205

Powell, A. (2005). Health care reform in China discussed. Harvard Gazette Arch. Retrieved from http://news.harvard.edu/gazette/2005/09.15/09-china.html

Qian, T., Zhang, L. M., & Li, Z. X. (2011). Algorithm of adaptive Fourier decomposition. *IEEE Transactions on Signal Processing*, *59*(12), 5899–5906. doi:10.1109/TSP.2011.2168520

Ren, Y., Pazzi, R. W. N., & Boukerche, A. (2010). Monitoring patients via a secure and mobile healthcare system. *IEEE Wireless Communications*, *17*(1), 59–65. doi:10.1109/MWC.2010.5416351

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sakoe, H., & Chiba, S. (1978). Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49. doi:10.1109/TASSP.1978.1163055

Sekar, B.D., Dong, M.C., Shi, J., & Hu, X.Y. (2012). Fused hierarchical neural network for cardiovascular disease diagnosis. *IEEE Sensors Journal*, *12*(3), 644–650. doi:10.1109/JSEN.2011.2129506

Sekar, B.D., Dou, J. Y., Fu, B. B., Fei, X. L., & Dong, M. C. (2011). Complexity and similarity analysis of heart sound for cardiovascular disease detection. *Science Alert Journals on Biotechnology*, *10*(4), 316–322.

Shi, J., Dong, M.C., Sekar, B.D., & Lei, W.K. (2009).Prognose coronary heart diseases through sphygmogram analysis and SVM classifier. Proceedings of the *7th International Conference on Information, Communications and Signal Processing*, Macau. IEEE. 10.1109/ICICS.2009.5397539

Shi, J., Sekar, B.D., Dong, M.C., & Lei, W.K. (2010). Fuzzy neural networks to detect cardiovascular diseases hierarchically. Proceedings of the *10th International Conference on Computer and Information Technology (CIT).*Bradford, UK. IEEE. 10.1109/CIT.2010.137

Song, T., Jamshidi, M. M., Lee, R. R., & Huang, M. (2007). A modified probabilistic neural network for partial volume segmentation in brain MR image. *IEEE Transactions on Neural Networks*, *18*(5), 1424–1432. doi:10.1109/TNN.2007.891635 PMID:18220190

Staszewski, W. J., & Robertson, A. N. (2007).Time–frequency and time–scale analyses for structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365*(1851), 449-477.

Tang, T. C. (2014). Noninvasive and invasive comprehensive intelligent cardiovascular disease diagnosis in *e*-home healthcare. *International Journal of Bioscience. Biochemistry and Bioinformatics*, *4*(5), 326–330.

Tomas, B., & Roncevic, Z. (2009). Spectral analysis of heart murmurs in children by Goertzel algorithm. Proceedings of *Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns (COMPUTATIONWORLD)*, Athens, Greece. IEEE. doi:10.1109/ComputationWorld.2009.25

University of Dundee. (n.d.). Human Heart Sounds and Murmurs. Retrieved from http://www.dundee.ac.uk/medther/Cardiology/hsmur.html

University of Michigan Heart Sound and Murmur Library. (n. d.). *Heart Sound Samples*. Retrieved from http://www.med.umich.edu/lrc/psb/heartsounds/index.htmhttp://www.med.umich.edu/lrc/psb/heartsounds/index.htm

University of Washington School of Medicine. (n.d.), Demonstrations - Heart Sounds & Murmurs Exam - Physical Diagnosis Skills -.*Heart Sound Samples*. Retrieved from http://depts.washington.edu/physdx/heart/demo.htmlhttp://depts.washington.edu/physdx/heart/demo.html

Ursino, M., Cavalcanti, S., Severi, S., & Sangiorgi, A. (1996). Mathematical analysis of the hemodynamic response during hemodialysis. *Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Amsterdam, the Netherlands: IEEE.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications*, *12*(2-3), 113–127. doi:10.100711036-007-0017-1

Vial, J., Noçairi, H., Sassiat, P., Mallipatu, S., Cognon, G., Thiébaut, D., & ... (2009). Combination of dynamic time warping and multivariate analysis for the comparison of comprehensive two-dimensional gas chromatograms: Application to plant extracts. *Journal of Chromatography. A*, *1216*(14), 2866–2872. doi:10.1016/j.chroma.2008.09.027 PMID:18834592

Wang, H., Peng, D., Wang, W., Sharif, H., Chen, H. H., & Khoynezhad, A. (2010). Resource-aware secure ECG healthcare monitoring through body sensor networks. *IEEE Wireless Communications*, *17*(1), 12–19. doi:10.1109/MWC.2010.5416345

White, C. C., Fang, D., Kim, E. H., Lober, W. B., & Kim, Y. (2006). Improving healthcare quality through distributed diagnosis and home healthcare (D2H2). Proceedings of the *1st Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare*, Virginia, USA. IEEE. 10.1109/DDHH.2006.1624823

World Health Organization. (2011). *Global Health and Aging*. Retrieved from www.who.int/topics/ageing/

Zhang, L., Ahn, G. J., & Chu, B. T. (2002).A role-based delegation framework for healthcare information systems. In *Proceedings of the seventh ACM symposium on Access control models and technologies*. Monterey, USA. 10.1145/507711.507731

Zhang, Z., Lall, C., & Chen, Y. (2011). Stability analysis of QRS features to evaluate signal quality for multi-lead QRS detection. Proceedings of *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Massachusetts, USA. IEEE.

Zhao, X.M., & Cao, G.T. (2009). A novel de-noising method for heart sound signal using improved thresholding function in wavelet domain. Proceedings of *International Conference on Future BioMedical Information Engineering (FBIE)*, Sanya, China. IEEE. 10.1109/FBIE.2009.5405795

Zheng, X.L., Fu, B.B., Fei, X.L., & Dong, M.C. (2013).Automatic pathological signal detection of heart sound based on multi-resolution analysis. Proceedings of the *10th International Conference on Mathematical Biology and Ecology (MABE)*, Milan, Italy.

KEY TERMS AND DEFINITIONS

Cardiovascular Disease: Refers to any pathological condition caused by the dysfunction of the cardiovascular system.

***e*-Health:** The healthcare practice supported by multiple key technologies including information and communication technology, mobile technology, body sensor networks, intelligent technology, etc.

Electrocardiogram: The biological signal which is used to measure/monitor the cardiac electrical activity through measuring the electric action potentials that are conducted through the heart muscle.

Embedded Link System: It is a computer system with hardware and software part specifically designed for a dedicated function, often used for real-time computing.

Heart Sound: The sound generated from the cardiac system due to mechanical activity of the heart, including blood flow, vibrations of chamber walls and opening and closing of valves.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Inference Engine: It is a software system that interprets and evaluates the rules/facts in the knowledge base defined based on expert's knowledge or large data analysis to draw the final conclusion.

Sphygmogram: It is the arterial pulse wave contour used to measure/monitor the pressure gradients caused when the heart contracts and the blood travels along the arterial walls.

Wireless Body Sensor Networks: Advances of wireless communication and network technologies linked with the computing and biological sensing techniques used towards the development of wireless network of wearable or implantable computing devices.

This research was previously published in Optimizing Assistive Technologies for Aging Populations edited by Yosry S. Morsi, Anupam Shukla, and Chandra Prakash Rathore; pages 207-250; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 38
Hybrid Ensemble Learning Methods for Classification of Microarray Data:
RotBagg Ensemble Based Classification

Sujata Dash
North Orissa University, India

ABSTRACT

Efficient classification and feature extraction techniques pave an effective way for diagnosing cancers from microarray datasets. It has been observed that the conventional classification techniques have major limitations in discriminating the genes accurately. However, such kind of problems can be addressed by an ensemble technique to a great extent. In this paper, a hybrid RotBagg ensemble framework has been proposed to address the problem specified above. This technique is an integration of Rotation Forest and Bagging ensemble which in turn preserves the basic characteristics of ensemble architecture i.e., diversity and accuracy. Three different feature selection techniques are employed to select subsets of genes to improve the effectiveness and generalization of the RotBagg ensemble. The efficiency is validated through five microarray datasets and also compared with the results of base learners. The experimental results show that the correlation based FRFR with PCA-based RotBagg ensemble form a highly efficient classification model.

INTRODUCTION

Cancer is caused due to the changes or mutation in the expression profiles of certain genes which elevates the importance of feature selection techniques to find relevant genes for classification of the disease. The most significant genes selected from the process are useful in clinical diagnosis for identifying disease profiles (Yang et al., 2006). The discriminative genes are selected through feature selection techniques that aim to select an optimal subset of genes. But, high dimension and small sample size characteristics of microarray dataset creates lot of computational challenges for selecting optimal subsets of genes

DOI: 10.4018/978-1-7998-1204-3.ch038

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

such as the problem of "curse of dimensionality" and over-fitting of training dataset. Feature selection is often used as a preprocessing step in machine learning. Only non-redundant and relevant features are sufficient enough to provide effective and efficient learning. However, selecting an optimal subset is very difficult (Kohavi & John, 1997) as the possible number of subsets grows exponentially when the dimension of the dataset increases.

The feature selection techniques can be broadly classified into filter (Hall, 2000; Liu, Motoda & Yu, 2002; Yu & Liu, 2003) and wrapper model (Hsu et al., 2011; Dash, Patra & Tripathy, 2012). The filter model uses specific evaluation criterion which is independent of learning algorithm to select feature subset from the dataset. It depends on various evaluation measures which are employed on the general characteristics of the training data such as information, distance, consistency and dependency. The wrapper method measures the goodness of the selected subsets using the predictive accuracy of the learning algorithm. But these methods require intensive computation for high dimensional dataset. Apart from this another key factor in feature selection is search strategy. The trade-off between optimal solution and computational efficiency is attained by adopting an appropriate search strategy such as random, exhaustive and heuristic search (Dash & Liu, 2003).

There are feature selection methods available for supervised (Yu & Liu, 2003; Dash & Liu, 1997) and unsupervised (Dash, Choi, Scheuermann & Liu., 2002) learning methods and it has been applied in several applications like genomic microarray data analysis, image retrieval, text categorization, intrusion detection etc. But, the theoretical and empirical analysis has demonstrated that the presence of irrelevant and redundant features (Kohavi & John, 1997; Hall, 2000) in the dataset reduces the speed and accuracy of the learning algorithms, thus need to be removed from the dataset. Most of the feature selection techniques employed so far has considered individual feature evaluation and feature subset evaluation (Guyon & Elisseeff, 2003; Abraham, 2004). Individual feature evaluation method ranks the features with respect to their capability of differentiating instances of different classes and eliminates the irrelevant and redundant features likely to have the same rankings. The feature subset evaluation method finds a subset of minimum features satisfying measure of goodness removes irrelevant and redundant features. It is observed that the advance search strategies like heuristic search and greedy search used for subset evaluation even after reducing the search space from $O(2^N)$ to $O(N^2)$ prove to be inefficient for high-dimensional dataset. This shortcoming encourages exploring different techniques for feature selection which will address both feature relevance and redundancy for high-dimensional microarray dataset.

There are various uncertainties associated with fabrication of microarray data such as the gathering of data, hybridization and image processing. They introduce lot of noises which need to be addressed by a robust and reliable classification model (Piao, Piao, Park & Ryu, 2012). On the other hand, conventional machine learning algorithms encounter many challenges to develop an effective and reliable classification model. The generalization capability of such kind of classifier algorithm based on a few significant genes and small number of training samples cannot be dependable. Therefore, it is essential to devise generalized robust classification methods which could overcome the constraints of small sample size and uncertainties associated with the datasets. This study motivates to develop ensemble classifier which is not very sensitive to the above specified constraints.

Ensemble learning technique is an advanced mechanism to combine multiple number of learning techniques to achieve better prediction accuracy (Liu et al., 2010; Dietterich, 2000; Yang, Yang, Zhou & Zomaya, 2010). It has the advantage of ignoring the constraint of sample size and the potential threat of overfitting by averaging and incorporating over multiple learning models (Hansen & Salamon, 1990). That is how the dataset is being used in an effective way, which was highly difficult for many bioin-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

formatics applications to handle. The research findings of ensemble learning have shown a promising improvement in the classification accuracy while dealing with data under uncertainties (Dietterich, 2000; Hansen & Salamon, 1990). However, the ensemble classifier only can outperform its own individual base learners if and only if its base learners are accurate and diverse (Dietterich, 2000). The base learners are considered as accurate, if the classification error rate of new samples is better than the randomly guessing classes. Similarly, two base learners can be regarded as diverse if their decisions on the same sample are different and always they do not converge to a single decision.

The ensemble methods employ a base learning algorithm to the training sets permutated differently. The popular examples of these techniques include Bagging (Dietterich, 2000; DeBock, Coussement & Vanden Poel, 2010), AdaBoost (Zhang & Zhang, 2010), Rotation Forest (Liu & Huang, 2008), Random Subspace (Breiman, 1996) and Random Forest (Zhang & Zhang, 2008). In bagging, bootstrap method is used to construct each base classifier from the original training set T i.e., a random sample is derived with replacement from the original dataset. Then majority voting is used to find the classification accuracy of the ensemble classifier where the new test data is being assigned with a class having highest number of votes obtained from the base learners (DeBock, Coussement & Vanden Poel, 2010).

Another ensemble method which is very popular in terms of simplicity and adaptability is AdaBoost (Zhang & Zhang, 2010). It builds an ensemble of classifiers by using a specified learning algorithm as base learner for the datasets that are obtained from the original training set either using resampling or reweighting (Zhang & Zhang, 2010). However, AdaBoost attempts to produce new "strong" classifiers that are able to better predict the hard instances of the training set.

Rotation Forest (Liu & Huang, 2008) is a variant of bagging ensemble method. But this ensemble classifier is built with a set of decision trees as the base learners. The bootstrap method is used to extract samples from the original training set to construct a new training set for each base learner. Subsequently, the features of the new training dataset are split randomly into some subsets, which are then transformed individually using some transformation techniques. As we know, a small perturbation may build a complete different tree but transformation techniques only can guarantee the diversity of the ensemble system (Liu & Huang, 2008; Dash, 2015). Moreover, at the same time Rotation Forest (Liu & Huang, 2008) enhances the generalization capability of the base learners as well as the diversity in the ensemble classifiers when compared with other ensemble classifiers, such as Bagging (Breiman, 1996), AdaBoost (Freund & Schapire, 1997), and Random Forest (Breiman, 2001).

C. Zhang and J. Zhang (2008) have developed a novel ensemble classifier "RotBoost" by combining two different ensemble techniques i.e., Rotation Forest and AdaBoost. They have replaced the base classifier of Rotation Forest with AdaBoost and established that the RotBoost performs better than the individual ensembles integrated together in this model. The performance of the model is also validated with gene related datasets from UCI repository.

In this paper, a hybrid ensemble learning methodology is proposed by combining a correlation based feature relevancy and redundancy (CO-FRFR) (Dash & Dash, 2014) feature selection method and PCA-based RotBagg ensemble classification technique. The efficiency of the model has been verified through publicly available microarray datasets and also compared with the experimental results of ICA-based-RotBoost (Osareh & Shadgar, 2013) technique. In addition to this, the efficiency of CO-FRFR-PCA- RotBagg has also been compared with the ensemble techniques used as base learners in the proposed model. To fix the values of the parameters of the base learners and of the ensemble model, a comparison among two variants of the RotBagg model such as PCA-based RotBagg and RP (Random Projection)-based RotBagg has been made. The integrated PCA-based RotBagg ensemble with CO-FRFR

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

exhibited superior average performance in terms of achieving generalization accuracy in comparison to other ensemble classifiers. Therefore, this model needs to be recommended as the most effective and efficient classification technique for predicting microarray datasets. The rest of the paper encompasses section 2 explaining feature relevance and redundancy concept and section 3 proposing the framework of RotBagg ensemble technique and feature selection method. In section 4 an empirical study of the proposed method in terms of efficiency and effectiveness comparing with other approaches are discussed. Section 5 provides a conclusion with future directions.

RELEVANCE AND REDUNDANCY OF FEATURES

It is observed that, a considerable number of features are not informative because of their irrelevance or redundancy (Abeel et al., 2010) to the class concept in high-dimensional dataset. Redundant and irrelevant features reduce classification accuracy and run-time of the system (Hall, 2000; Dash, Choi, Scheuermann & Liu., 2002). Conventional approaches of feature selection have given more emphasis on relevant data but very little attention to feature redundancy (Kohavi & John, 1997; Liu et al., 2010). The features of the dataset can be broadly classified into three categories considering feature relevance characteristic namely strongly relevant, weakly relevant and irrelevant (Kohavi & John, 1997). These three concepts of feature relevance cannot handle redundancy all alone. The formal definition can be stated as: Let F be a full set of features and for a feature F_i, $S_i = F - \{F_i\}$

Definition 1 (Strong relevance) *A feature Fi is strongly relevant iff* $P(C\backslash Fi,Si) \neq P(C\backslash Si)$.

Definition 2 (Weak relevance) *A feature Fi is weakly relevant iff* $P(C|Fi,Si) = P(C|Si)$ and $\exists S_i' \subset S_i$, *such that* $P(C \mid F_i, S_i') \neq P(C \mid S_i')$

Definition 3 (Irrelevance) *A feature Fi is irrelevant iff for all* $S_i' \subseteq S_i$, $P(C \mid F_i, S_i') = P(C \mid S_i')$

An optimal subset contains strong relevant features and the removal of which always affects the conditional class distribution. In case of weak relevance, the feature may not always require for an optimal subset but at certain conditions it may be required. Irrelevance suggests that the feature is not required. An optimal subset should contain all strongly relevant features, none of the irrelevant features, and a subset of weakly relevant features. Unfortunately, the definitions do not explain which of the weakly relevant features need to be selected for the optimal subset and which are not. Therefore, at this stage feature redundancy need to be investigated among relevant features. It is observed that, two features are redundant to each other if they are correlated completely. But, when a feature is partially correlated with a set of features, it may not be so easy to determine the feature redundancy existing in the dataset.

Correlation Measures

In feature selection process, we deal with mainly two types of correlation measures namely, linear and non-linear which exist between genes or between target class and genes. In the proposed method, non-linear correlation measures which are based on the information –theoretical concept of *entropy,* a measure of the uncertainty of a random variable is applied. The entropy of a variable X is defined as:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$H(X) = -\sum iP(x_i)\log_2(P(X_i))$$

and the entropy of *X* after observing values of another variable *Y* is defined as

$$H(X / Y) = -\sum jP(y_j)\sum iP(x_i / y_j)\log_2(P(x_i / y_j))$$

where the prior probability is $P(x_i)$ for all X_i and the posterior probability for all X given the values of Y is $P(x_i/y_i)$. The additional information given by Y to X is called *information gain,*

$$IG(X / Y) = H(X) - H(X / Y)$$

In this article, symmetrical uncertainty has been used to make up for information gain's bias towards high valued features and confines the values in the range of [0, 1]. Value 1 informs that if we know the value of either of the features we can completely predict the value of the other and value 0 indicates both are independent of each other. It can be expressed as:

$$SU\left(X,Y\right) = 2\left[\frac{IG(X / Y)}{H(X) + H(Y)}\right]$$

ENSEMBLE FRAMEWORK AND FEATURE SLECTION METHOD

RotBagg Ensemble Framework

The proposed ensemble method, RotBagg, is a hybrid intelligent classifier developed by combining the concept of Rotation Forest and Bagging ensemble classifier techniques with the objective of achieving higher accuracy than either of the individual techniques. Rotation Forest is a robust multiple classifier system which employs base classifiers which are accurate and the diversity among them is low. The framework of this method requires a linear transformation technique to project the dataset of each classifier or base learner into a new feature space and then is described as follows. In the framework of Rotation Forest, a linear transformation method is required to project data into new feature space for each classifier, and then the base classifiers are trained in different spaces so as to increase the accuracies of base classifiers and reduce the diversity in the ensemble system. Principal component analysis (PCA), non-parametric discriminant analysis (NDA) and random projections (RP) are used as feature transformation techniques in Rotation Forest.

Rotation Forest is a multiple classifier system which extracts different set of features to train independently all the *L* decision trees (base learners) (Breiman, (2001)). Let X be the training set, ω be the labels of each training sample and F *i*s the feature set. Assuming that there are N *t*raining samples with n *f*eatures in a microarray dataset, then X *i*s an N×*n m*atrix. Let ω be the set of class labels $\{\omega 1, \omega_2$. . . ω m} ass$_i$gned to each training sample. The training of base classifiers Ci req*u*$_i$res a mapping to be established between the input sample spaces and the class labels. If the feature set F is s*p*lit randomly

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

into k subsets of approximate size, then there may be L base classifiers required for the Rotation Forest which can be denoted by D1, . . ., D*L. There*fore, L and *k* are *t*wo important parameters which need to be determined carefully.

Now the following steps need to be followed to process the training dataset of an individual classifier Di.

Step 1: Randomly split F into k disjoint subsets and assume that the size of each subset be M = *n/k.*

Step 2: Consider F*ij* be the j th subset of features for training classifier D*i,* and X*ij* be the dataset X for the features in F*ij.*

Step 3: Select randomly a nonempty subset of classes from *Xij* for each subset. Then apply bootstrap technique to extract 75% of data from the original dataset to form a new training dataset, which is denoted by *Xij*'.

Step 4: Apply a linear transformation technique on *Xij*' and generate coefficients in matrix *Qij*. Denote the coefficients by *aij*1, ⋯, *aij*mj. This indicates the size of Xij' i*s M* x 1.

Step 5: At the end, construct a sparse rotation matrix Ri u*si*ng the coefficients generated in matrix Qij *(C.*-X. Zhang and J.-S. Zhang, 2010).

$$R_i = \begin{pmatrix} a_{i1}^{(1)},\ldots,a_{i1}^{(M_1)} & \{0\} & \{0\} \\ \{0\} & a_{i2}^{(1)},\ldots..,a_{i2}^{(M_2)} & \{0\} \\ \{0\} & \{0\} & a_{ik}^{(1)},\ldots...,a_{ik}^{(M_k)} \end{pmatrix}$$

Now the columns of sparse rotation matrix are rearranged as per the original feature sequence and it is denoted by Rai. Then the transformed training dataset for the corresponding classifier Di would be represented by XRai. The above process shows all the classifiers will be trained in a parallel manner.

During the validation phase of the classification process, for a test sample x, the classifier Di will have the probability dij (XRai) for the hypothesis that x belongs to the class φj . The confidence of a class can be computed by using average combination method which is shown below:

$$\varphi_j(x) = \frac{1}{L}\sum_{i}^{L} = d_{ij}(xR_i^a), \text{j} = 1,\ldots,\text{m}$$

In the proposed RotBagg ensemble classifier, the base classifiers D1, . . ., DL of Rotation Forest are replaced by Bagging classifiers. Bagging (Bootstrap aggregating) was introduced by Breimen (Breiman L., 1996) and it randomly replaces the original training dataset by N items. The replaced training sets are known as bootstrap replicates wherein some instances may not appear but others may appear more than once. The ensemble classifier is constructed by aggregating the base classifiers where all of them have equal vote.

The integration of bagging with rotation forest enhances the computational ability in terms of accuracy and diversity of the integrated RotBagg ensemble classifier. RotBagg is more advantageous than bagging because it has the ability of parallel execution.

The pseudocode 1 explains how each sub-ensemble created by bagging learns without depending on other ensembles. The learning principle is explained through pseudocode 1. Some of the important

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

parameters of RotBagg ensemble classifier need to be decided beforehand i.e., the values of S and T which specify the number of iterations performed by rotation forest and bagging respectively. Similarly, the values of another two parameters such as k and M which specify the number of subsets and the size of each subset respectively need to be fixed. After fixing the above parameters, the RotBagg ensemble algorithm can be implemented. However, the fast decision tree learners (REP Tree) are being employed as the base learners of the proposed RotBagg ensemble algorithm.

Transformation Techniques

As it is known that decision trees are used as the base learners to construct rotation forest and RotBagg, they need to satisfy the two characteristics of ensemble system i.e., accuracy and diversity. A bootstrap mechanism is used to form a new training set from the original for each base classifier of the ensemble. Then the new training set is randomly split into disjoint subsets which are then transformed individually by a linear transformation technique. The transformed features of each tree, subsequently reconstruct the full feature set of the ensemble. Since a small rotation in the axis creates a different tree, the transformation technique can guarantee the quality of diversity of the ensemble system. As a matter of fact, all the outputs are fused by the average rule. Mostly, there have been three types of transformation techniques adopted by different researchers (Kuncheva & Rodriguez, 2007) such as non-parametric discriminant analysis (NDA), principal component analysis (PCA) and random projections (RP). It has been found that PCA -based rotation forest performs better than the other two. Indeed, the principal components use to preserve all the discriminatory information of the dataset after transformation and the rotation of axes primarily generate diversity among the base classifiers instead of weakening the individual classifiers. In addition to this, an efficient sparse random projection (RP) method has been mentioned (Kuncheva & Rodriguez, 2007) with effective performance. Therefore, in this paper PCA and RP method has also been considered for comparison in the microarray dataset classification experiment.

Pseudocode-1 of RotBagg Ensemble Classifier

Input

1. **L**: a training set, L= $\{(x_i, y_i)\}_{i=1}^{N}$ = [XY]

where X is an N × n matrix containing the input values and y is an N –dimensional column vector containing the class labels.

2. **K**: number of attributes subsets (or M: number of input attributes contained in each subset).
3. **C**: a base learner.
4. **S**: number of iteration for Rotation Forest.
5. **T**: No of iteration for Bagging.
6. **x**: a data point to be classified.

Training Phase
For s = 1, 2, ..., S

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. Use the steps similar to those in Rotation Forest to compute the Rotation matrix, say, R_s^a and let $L^a = [XR_t^aY]$ be the training set for classifier C_5.
2. Initialize the weight distribution over L^a as $D_1(i) = 1/N(i=1,2,\ldots,N)$
3. For t = 1, ..., T
   a. According to distribution D_t perform N extractions randomly from L^a with replacement to compose a new set L_t^a
   b. Apply ω to L_t^a to train a classifier C_t^a and then compute the error of C_t^a as

$$\varepsilon_t = Pr_{i\sim D_t}(C_t^a(x_i) \neq y_i) = \sum_{i=1}^{N}(C_t^a(x_i) \neq y_i))D_t(i)$$

   c. If $\varepsilon_t > 0.5$ then set D_t (i) = 1/N (i = 1, 2, ..., N) and go to step(a); if $\varepsilon_t = 0$, then set $\varepsilon_t = 10^{-10}$ to continue the following iterations
   d. Choose $\alpha_t = 1/2\text{In}((1- \varepsilon_t)/ \varepsilon_t)$
   e. Update the distribution D_t over L^a as:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t}, if C_t^a(x_i)=y_i \\ e^{\alpha_t}, if \ C_t^a(x_i)=y_i \end{cases}$$

where Z_t is a normalization factor being chosen so that D_{t+1} is a probability distribution over L^a .

End for

4. Let C_s (x) = argma $x_{y\in 0}\sum_{t=1}^{T}\alpha_t I(C_t^{\alpha}(x) = y)$

End for
Output

5. The class label for x predicted by the final ensemble C^* as

$$C^*(x) = \text{argma } x_{y\in 0}\sum_{s=1}^{s} I(C_s(x) = y)$$

where I (.) is an indicator function.

Feature Selection

The sample size of microarray datasets available in public database is very small compared to the number of attributes involved and this ratio problem is referred to as curse of dimensionality. This problem becomes a challenge for many classification techniques. As a matter of fact, only a small subset of attributes of microarray dataset is responsible for causing cancer. This necessitates the development of a data reduction technique to remove irrelevant, redundant and noisy data to find an effective solution

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to increase the classification accuracy. On the other hand, if RotBagg ensemble is applied directly to the dataset without reduction then a rotation matrix with thousands of dimensions is required for each tree which increases the computational complexity to a large extent. There are two major categories of feature selection algorithms (Dash & Dash, 2014) such as filter and wrapper algorithm used to find the predictive genes. The filter model finds the significant attributes using only the general characteristics of the dataset without the learning algorithm involvement whereas the wrapper model uses incremental approach involving an evaluation function to validate the subsets of features generated. Although, this method finds subset of features which provides a superior learning performance, still they are computationally expensive than filter model (Dash, Patra & Tripathy, 2012). In this paper, a correlation based feature relevance and feature redundant (CO-FRFR) technique which is already applied in (Dash & Dash, 2014) and suggest that CO-FRFR identifies the redundancy among relevant features. The advantage of this model is it uses interdependence of features along with the dependence to the class and identifies predominant features by reducing the dimension to a greater extent. Thus, it improves the classification accuracy in return. It has been explained in pseudocode-2.

Pseudocode-2 of Correlation Based Feature Relevancy and Redundancy Algorithm (CO-FEFR)

Relevance Analysis

1. Rank the features according to the decreasing values of $SU_{i,c} > Rel\text{-}parameter$
2. Redundancy analysis
3. Initialize F_i with the value of first feature in the list
4. Compare and eliminate features for which F_i forms an approximate redundant cover
5. Set F_i as the predominant feature in the list
6. Repeat step 3 until the end of the list
7. Evaluate the predictive significance of the predominant feature list using RotBagg ensemble

EXPERIMENTAL STUDY

Microarray Dataset

The efficiency of the proposed method is measured through three publicly available multi-category and two binary microarray datasets. The datasets include MLL Leukemia, Lung cancer, ovarian cancer, CNS and SRBCT. Table 1 summarizes the characteristics of the microarray datasets. The out-liars of the datasets are removed through preprocessing which is an important step for handling gene expression data. This includes two steps: filling missing values and standardization. The missing values of the datasets are filled using the average value of the corresponding gene. Standardization is performed so that every gene expression has mean 0 and standard deviation 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Microarray Data sets used in this experiment

DATA SET	No. of Features	No. of Samples	No. of Classes	References
MLL Leukemia	12534	72	03	Armstrong et al., 2002
Ovarian Cancer	15154	253	02	Emanuel F Petricoin III, et al, 2002
Lung Tumour	12601	203	05	Arindam Bhattacharjee, et al., 2001
SRBCT	2308	83	04	Khan J et al., 2001
CNS	7129	60	02	Scott L. Pomeroy, et al. 2002

Experimental Design and Analysis

In this work, we have used open source tool WEKA to implement the proposed RotBagg ensemble classifier and the feature selection / reduction techniques. The efficiency of the proposed feature reduction technique (CO-FRFR) has been compared with two representative feature reduction methods such as correlation-based feature selection (CFS) and ReliefF in this experiment. The dataset is cleaned removing all missing values and converting it to mean 0 and standard deviation 1.

ReliefF (Dash & Dash, 2014; Ambroise & McLachlan, 2002) has the ability to deal with multiclass problems and more robust towards noisy and incomplete data. It evaluates the worth of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different class. The CFS method (Hall, 2000) depends on a set of heuristics to assess the adequacy of the subsets. The heuristics used in this method take in to consideration both the correlation among the features and the dependency of features to the class.

To select the most informative features and reduce the computational complexity of the problem, all the feature selection method such as CFS, ReliefF and CO-FRFR are used to derive the features for each method. The selected features are shown in Table 2 for each datasets and it is evident from the table that the selection of features is different for each dataset and largely depends on the feature selection algorithms. It is quite apparent from Table 2 that CO-FRFR selects least number of predictive features while achieving highest reduction in dimension of the datasets. Therefore, it conforms to the theoretical concept of CO-FRFR to recognize and ignore redundant features.

The classification accuracy of the selected features is evaluated using RotBagg ensemble using a fast decision tree (REP tree) as the base learner and compared with the performance of integrated ensembles RotBagg and also with the base learner performance which is shown in the Tables 3 to 14. This learner is applied to the reduced datasets obtained from the feature selection algorithms and the overall classification accuracy is recorded using stratified 10-fold cross validation method to achieve statistically reliable predictive measurements (Dash & Liu, 2003; Dash & Liu, 1997). It is observed from the table that, CO-FRFR outperforms the other two feature selection algorithms in terms of achieving least number of features. The classification accuracy, precision, F-score, MCC, ROC area of all the datasets are significantly better for the CO-FRFR based hybrid ensemble RotBagg when compared with other combination of feature selection methods. Two types of transformation techniques such as PCA and RP are used along with the hybrid ensemble classification. It is observed from the experiment that PCA based transformation outperforms RP transformation in hybrid RotBagg as well as in single Rotation Forest ensemble. To compare the classification performance of the classifiers, 10-fold cross validation has been used for all the experiments. In this method, each dataset is divided into 10 disjoint sets, each

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

having roughly same distribution. But, the single classifier such as a decision tree does not provide results like an accurate classifier when it is applied to microarray gene classification problem which suffers lot of limitations like uncertainties, noises and curse of dimensionality. The results obtained from Table 6 to Table 14 conforms the findings of (Osareh & Shadgar, 2013).

The challenges encountered by a single decision tree algorithm have been overcome by a robust and accurate ensemble algorithm "RotBagg". This ensemble classifier operates on the discriminative features obtained from the feature selection method CO-FRFR, ReliefF and CFS. Along with the RotBagg, another two ensemble classifiers namely rotation forest and bagging are also developed and trained using the reduced data in this experiment. All the ensembles have used a decision tree as a base learner (Yang, Yang, Zhou & Zomaya, 2010) because it is more sensitive to the changes made in the dataset. A first decision tree learner (REP tree) was used by all the methods as the base learners without pruning. Now each of the ensemble methods will train a large number of base classifiers nearly 100 in numbers to construct the ensemble classifiers. The total number of iterations for both rotation forest and bagging can be balanced by setting $S = T = 10$ which is similar to (Osareh & Shadgar, 2013). As per the concept, a moderate value needs to be set for the parameter M according to the size of the feature set F. Therefore, the performance of RotBagg will vary for different size of the features contained in each subset M. But, when compared the performance of the classification results varying M from 1 to 20, it is observed that the variation is insignificant. This shows there is no consistent relationship exists between M and the learning accuracy (Ambroise & McLachlan, 2002). In this experiment an optimum value is set for M i.e., $M = 3$ to established a consistent relationship between the diversity and classification accuracy of the base learners.

To find the most effective transformation technique to maintain the diversity and accuracy of the base classifiers, two well-known transformation techniques i.e., PCA and RP have been used in this experiment. After the transformation of the datasets, the principal components preserve the discriminatory information. Then the axes are rotated optimally which primarily develops the diversity among the classifiers rather than weakening the individual classifiers. As a result, a trade-off between diversity and accuracy can be maintained.

The small number of instances in the microarray dataset led to unreliable solutions. To counter this problem, 10-fold cross validation method (Peng, 2006) was suggested to split the datasets into 10 equal fold to validate the solution. However, in this work, a 10-fold cross validation is deployed to evaluate the performance of the feature selection algorithm.

Table 3 shows CO-FRFR based RotBagg classification accuracy, precision, F-score, MCC, ROC area and FP rate for all the five datasets when both the transformation matrix are applied. The classifier is trained 10 times such that in each iteration a different set is kept for testing. The AUROC (Area under the Receiver Operating Characteristic) value and MCC are computed as the average of each of these iterations. Many statistical techniques have been applied to compare the performance of classifier across multiple datasets. Again the performance of the classifiers can be ranked (Kuncheva & Whitaker, 2003) using average AUROC values with 1 being the best. Similarly, Friedman and Bonferroni-Dunn tests (Kuncheva & Whitaker, 2003) are performed to determine whether or not the newly introduced methods are statistically better than the existing methods. It is quite apparent from Table 3 that PCA-based RotBagg performs significantly better than RP-based RotBagg for all the datasets and achieves full rank for ovarian cancer and SRBCT. It can also be noted from Table 3 that PCA–based RotBagg ensembles outperform RP-based counterparts in terms of accuracy of classification. Also it is important to record that all these experiments have been performed on the features selected by CO-FRFR algorithm.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

After choosing the transformation method, a comparative study on the performance of PCA/RP-based RotBagg using ReliefF and CFS feature selection methods to reduce the feature space is made. Table 4 and 5 demonstrates the classification accuracy of each classification method on the selected datasets. CO-FRFR based PCA-RotBagg ensemble performs significantly better than bagging and rotation forest invariably for all the datasets.

The most important issue of ensemble learning technique is the balance between diversity and accuracy of the classifiers involved in the ensemble as base learners. The base learners in an ensemble should have high accuracy and should avoid making misclassification errors. Thus, the accuracy of fused outputs will be more than that of the individual classifier (Ambroise & McLachlan, 2002). However, the diversity conflict with accuracy of base learners, i.e., the more accurate the base learners are, the lower the diversity is. So, it is difficult to define a single measure of diversity that measures the performance of ensemble. In this experiment, pair wise diversity measure has been utilized to measure the ability of the proposed PCA-based RotBagg which can build accurate and diverse base learners effectively.

Comparing the performance of Table 3, 4 and 5 with 6, 7 and 8, it is understood that the performance of hybrid ensemble is more significant when paired with CO-FRFR. But it is also apparent from the table that irrespective of the feature selection method chosen the performance of the hybrid RotBagg is most effective than the performance of the individual ensembles and the decision tree learner.

The classification performance of the RotBagg ensemble obtained for the reduced dataset derived from ReliefF and CFS is compared with the performance of individual ensemble and observed that RotBagg performs better than others can be observed from Table 9, 10, 11, 12, 13 and 14.

In order to measure the level of agreement between a pair of base learners, a k-metric or kappa statistic (Pomeroy et al, 2002) is used while correcting the chance. For c- class labels, k is defined on $c \; x \; c$ coincidence matrix M of the two classifiers. The entry $m \; \mathrm{k}_{,s}$ of M is the proportion of the data set, which Di labels as wk and Dj labels as ws. Now, the agreement between Di and Dj is given by

$$\mathrm{Ki,j} = \frac{\sum Km_{k,k} - \mathrm{AB}C}{1 - \mathrm{AB}C}$$

If_{th}e value of $k = 1$ then the classifiers decisions are the same otherwise $k = 0$, which represent the case when the classifiers are entirely independent, and the agreement of the two classifiers equals that expected by chance (Cohen, 1960). Table 3 shows the value of kappa statistic k for both P*C*A-based-RotBagg and RP-based- RotBagg ensemble classifiers. It is observed from the table that for ovarian cancer and SRBCT the value of k is 1 i.e., the classifiers decisions are same whereas the value of k for CNS is 0.65 in PCA-based RotBagg. In Table 3, 4 and 5, the value of kappa statistic is very close to 1 for RotBagg ensemble for all the three feature selection methods whereas it is significantly less for all the individual ensembles shown in remaining tables.

CONCLUSION

In this study, we proposed a novel concept of correlation based feature relevancy and redundancy (CO-FRFR) selection algorithm and a hybrid robust ensemble classifier, "RotBagg". This ensemble technique preserves the most essential characteristic of an ensemble framework, i.e., accuracy and diversity. In

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. No. of selected genes for each gene selection algorithms

Dataset	Initial Attribute/Sample nos	Co-FRFR	ReliefF	CFS
MLL	12533/72	23	34	150
Ovarian cancer	15154/253	30	42	36
Lung	12600/203	15	28	550
SRBCT	2308/83	14	38	119
CNS	7130/60	14	29	40

Table 3. Stratified 10-fold-CV learning performance of CO-FRFR based RotBagg ensemble using PCA/ RP transformation methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	97.22/ 95.833	0.975/ 0.959	0.972/ 0.958	0.960/ 0.938	**0.999/ 0.996**	0.9581/ 0.9371	0.011/ 0.020
Ovarian	*99.2095/ 98.8142*	0.992/ 0.988	0.992/ 0.988	0.983/ 0.974	**1.0/ 0.999**	0.9828 /0.9741	0.014/ 0.021
Lung	93. 1034 / 92.1182	0.932 / 0.924	0.927/ 0.914	0.849/ 0.832	**0.986/ 0.975**	0.8544/ 0.8304	0.129/ 0.151
SRBCT	**100 / 97. 5904**	1.0 / 0.976	1.0 / 0.976	1.0 / 0.967	**1.0 / 0.998**	1.0 / 0.9665	0.0 / 0.012
CNS	85.00/ 83.333	0.853 / 0.849	0.844 / 0.821	0.664 / 0.632	**0.912 / 0.897**	0.6512/ 0.5912	0.235/ 0.288

Table 4. Stratified 10-fold-CV learning performance of ReliefF based RotBagg ensemble using PCA/RP Transformation methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	97.22/ 95.833	0.975/ 0.959	0.972/ 0.958	0.960/ 0.938	**0.996/ 0.998**	0.9581/ 0.9371	0.011/ 0.020
Ovarian	99.2095/ 98.8142	0.992/ 0.988	0.992/ 0.988	0.983/ 0.974	**1.0/ 0.999**	0.9828 /0.9741	0.014/ 0.021
Lung	93. 1034 / 92.1182	0.932 / 0.924	0.927/ 0.914	0.849/ 0.832	**0.986/ 0.975**	0.8544/ 0.8304	0.129/ 0.151
SRBCT	100 / 97. 5904	1.0 / 0.976	1.0 / 0.976	1.0 / 0.967	**1.0 / 0.998**	1.0 / 0.9665	0.0 / 0.012
CNS	85.00/ 83.333	0.853 / 0.849	0.844 / 0.821	0.664 / 0.632	**0.912 / 0.897**	0.6512/ 0.5912	0.235/ 0.288

Table 5. Stratified 10-fold-CV learning performance of Correlation Feature Selection (CFS) based RotBagg ensemble using PCA/RP Transformation methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	100/ 98.6111	1.0/ 0.987	1.0/ 986	1.0/ 0.979	**1.0/ 1.0**	1.0/ 0.9789	0.0/ 0.009
Ovarian	99.6047/ 97. 6285	0.996/ 0.976	0.996/ 0.976	0.991/ 0.948	**1.0/ 0.999**	0.9914/ 0.9483	0.007/ 0.033
Lung	93. 1034/ 91. 133	0.936/ 0.918	0.916/ 0.898	0.839/ 0.814	**0.978/ 0.976**	0.8466/ 0.8382	0.110/ 0.120
SRBCT	98.7952/ 97.5904	0.988/ 0.977	0.988/ 0.976	0.982/ 0.966	**1.0 / 1.0**	0.9833/ 0.9667	0.005/ 0.009
CNS	81.6667/ 81.6667	0.816/ 0.823	0.810/ 0.806	0.585/ 0.587	**0.895/ 0.889**	0.5736/ 0.5635	0.275/ 0.297

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Stratified 10-fold-CV learning performance of CO-FRFR based Rotation Forest ensemble using PCA/RP Transformation methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	95.333/ 94.444	0.959/ 0.948	0.958/ 0.945	0.937/ 0.919	**0.999/ 0.993**	0.9368/ 0.9162	0.023/ 0.025
Ovarian	98.8142/ 97.6285	0.988/ 0.976	0.988/ 0.976	0.974/ 0.948	**0.999/ 0.998**	0.9741/ 0.9483	0.021/ 0.033
Lung	90.6404/ 90.6404	0.909/ 0.904	0.899/ 0.896	0.798/ 0.987	**0.984/ 0.983**	0.7997/ 0.7972	0.173/ 0.163
SRBCT	97.5904/ 96.3855	0.976/ 0.965	0.976/ 0.964	0.966/ 0.948	**0.998/ 0.995**	0.9666/ 0.9499	0.012/ 0.017
CNS	86.6667/ 83.3333	0.868/ 0.849	0.863/ 0.821	0.702/ 0.632	**0.920/ 0.910**	0.6935/ 0.5984	0.204/ 0.288

Table 7. Stratified 10-fold-CV learning performance of CO-FRFR based Bagging ensemble for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	91.6667	0.916	0.916	0.875	**0.984**	0.874	0.041
Ovarian	97.2332	0.972	0.972	0.940	**0.993**	0.9395	0.040
Lung	93.5961	0.937	0.933	0.871	**0.987**	0.8656	0.108
SRBCT	97.5904	0.976	0.976	0.976	**0.998**	0.9665	0.012
CNS	80.00	0.808	0.786	0.547	**0.897**	0.5181	0.327

Table 8. Stratified 10-fold-CV learning performance of CO-FRFR based first decision tree (REP Tree) learner for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	91.6667	0.920	0.917	0.876	**0.929**	0.8744	0.039
Ovarian	97.2332	0.972	0.972	0.940	**0.965**	0.9398	0.035
Lung	91.133	0.914	0.912	0.821	**0.916**	0.8216	0.100
SRBCT	80.7229	0.826	0.810	0.739	**0.896**	0.7332	0.077
CNS	75.00	0.743	0.741	0.427	**0.696**	0.4186	0.354

Table 9. Stratified 10-fold-CV learning performance of ReliefF based Rotation Forest ensemble using PCA/RP Transformation methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	94.4444/ 93.0556	0.946/ 0.932	0.945/ 0.930	0.916/ 0.896	**0.990/ 0.986**	0.9156/ 0.8944	0.032/ 0.039
Ovarian	99.6047/ 98.8142	0.996/ 0.998	0.996/ 0.988	0.991/ 0.974	**1.0/ 0.994**	0.9914/ 0.9742	0.007/ 0.016
Lung	91.1333/ 90.6404	0.911/ 0.907	0.908/ 0.899	0.812/ 0.798	**0.970/ 0.958**	0.818/ 0.7996	0.121/ 0.163
SRBCT	97.5904/ 95.1807	0.976/ 0.952	0.976/ 0.952	0.976/ 0.952	**0.998/ 0.996**	0.9667/ 0.9332	0.010/ 0.025
CNS	80.00/ 75.00	0.822/ 0.761	0.780/ 0.721	0.555/ 0.422	**0.806/ 0.772**	0.5062/ 0.375	0.349/ 0.420

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. Stratified 10-fold-CV learning performance of ReliefF based Bagging ensemble methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	91.6667	0.916	0.916	0.875	**0.984**	0.875	0.041
Ovarian	97. 2332	0.976	0.976	0.948	**0.995**	0.9483	0.033
Lung	91.6256	0.916	0.915	0.818	**0.963**	0.8284	0.121
SRBCT	92.7711	0.932	0.929	0.903	**0.993**	0.9005	0.024
CNS	73.33	0.725	0.726	0.392	**0.802**	0.387	0.363

Table 11. Stratified 10-fold-CV learning performance of ReliefF based first decision tree (REP Tree) learner for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	91.6667	0.917	0.915	0.877	**0.960**	0.8734	0.042
Ovarian	96.8379	0.968	0.968	0.931	**0.950**	0.931	0.042
Lung	85.2217	0.852	0.853	0.716	**0.890**	0.7116	0.126
SRBCT	85.5422	0.858	0.855	0.857	**0.924**	0.8006	0.051
CNS	70.00	0.694	0.700	0.696	**0.700**	0.3258	0.381

Table 12. Stratified 10-fold-CV learning performance of CFS based Rotation Forest ensemble using PCA/RP Transformation methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	98.6111/ 95.8333	0.987/ 0.959	0.986/ 0.958	0.986/ 0.958	**0.999/ 0.995**	0.9789/ 0.9371	0.009/ 0.020
Ovarian	99.6047/ 97.2332	0.996/ 0.972	0.996/ 0.972	0.991/ 0.940	**1.0/ 0.999**	0.9914/ 0.9395	0.002/ 0.040
Lung	91.6256/ 91.133	0.916/ 0.911	0.906/ 0.895	0.830/ 0.810	**0.989/ 0.989**	0.8227/ 0.8086	0.131/ 0.152
SRBCT	97.5904/ 96.3855	0.976/ 0.964	0.976/ 0.964	0.966/ 0.948	**0.999/ 0.995**	0.9667/ 0.9499	0.010/ 0.017
CNS	80.00/ 75.00	0.808/ 0.781	0.800/ 0.750	0.786/ 0.712	**0.906/ 0.870**	0.5181/ 0.359	0.327/ 0.442

Table 13. Stratified 10-fold-CV learning performance of CFS based Bagtging ensemble methods for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	91.6667	0.917	0.914	0.878	**0.983**	0.8726	0.020
Ovarian	97.6285	0.976	0.976	0.948	**0.989**	0.9483	0.033
Lung	91.133	0.911	0.897	0.813	**0.984**	0.8087	0.151
SRBCT	95.1807	0.952	0.952	0.934	**0.995**	0.9336	0.016
CNS	78.3333	0.783	0.758	0.516	**0.794**	0.4583	0.380

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 14. Stratified 10-fold-CV learning performance of CFS based first decision tree (REP Tree) learner for all 5 datasets

Datasets	% of Accuracy	Precision	F-score	MCC	ROC Area	Kappa Statistic	FP Rate
MLL	80.5556	0.807	0.806	0.708	**0.896**	0.7053	0.099
Ovarian	96.4427	0.964	0.964	0.923	**0.958**	0.9226	0.044
Lung	87.6847	0.884	0.887	0.871	**0.920**	0.7542	0.105
SRBCT	80.7229	0.818	0.807	0.807	**0.879**	0.7333	0.076
CNS	61.6667	0.596	0.617	0.602	**0.582**	0.1085	0.514

order to overcome the limitation of curse of dimensionality, uncertainties and noises, a small subset of most informative features were selected using 3 representative feature selection techniques namely, CFS, CO-FRFR and ReliefF. Then the ensemble classifier RotBagg was employed on the selected datasets. The experiment was conducted using two types of transformation matrices, i.e., PCA and RP method to implement RotBagg ensemble.

The efficiency of the proposed ensemble was assessed with different ensemble algorithms including rotation forest and bagging. The experimental results showed that the combination of CO-FRFR feature selection and PCA-based RotBagg ensemble with decision tree as base learners construct a robust classifier for the microarray datasets. The proposed method has achieved the highest generalization ability in comparison to its counterparts. It has also established a trade-off between the diversity and accuracy of the base learners involved.

REFERENCES

Abraham, A. (2004). Meta-Learning Evolutionary Artificial Neural Networks. *Neurocomputing*, *56*, 1–38. doi:10.1016/S0925-2312(03)00369-2

Ambroise, C., & McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(10), 6562–6566. doi:10.1073/pnas.102102699 PMID:11983868

Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., Minden, M. D., ... Korsmeyer, S. J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics*, *30*(1), 41–47. doi:10.1038/ng765 PMID:11731795

Bhattacharjee, A., Richards, W. G., Staunton, J., Li, C., Monti, S., Vasa, P., ... Meyerson, M. (2001). Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *PNAS. The National Academy of Sciences, USA*, *98*(24), 13790–13795. doi:10.1073/pnas.191502998 PMID:11707567

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. doi:10.1007/BF00058655

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi:10.1023/A:1010933404324

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. doi:10.1177/001316446002000104

Dash, M., Choi, K., Scheuermann, P., & Liu, H. (2002). Feature selection for clustering – a filter solution. *Proc. of the Second International Conference on Data Mining* (pp. 115–122).

Dash, M., & Liu, H. (1997). Feature selection for classification, *Intelligent Data Analysis*. *International Journal (Toronto, Ont.)*, *1*(3), 131–156.

Dash, M., & Liu, H. (2003). Consistency-based search in feature selection. *Artificial Intelligence*, *151*(1-2), 155–176. doi:10.1016/S0004-3702(03)00079-1

Dash, S. (2015). A Diverse Meta learning ensemble technique to handle imbalanced microarray dataset. *Proceedings of Seventh World Congress on Nature and Biologically Inspired Computing (NaBIC2015), Advances in Nature and Biologically Inspired Computing, Advances in Intelligent Systems and Computing*. DOI:10.1007/978-3-319-27400-3_1

Dash, S., & Dash, A. (2014). A correlation based multilayer perception algorithm for cancer classification with gene-expression dataset. *Proceedings of the International conference on Hybrid Intelligent Systems (HIS)*. 10.1109/HIS.2014.7086190

Dash, S., Patra, B., & Tripathy, B. K. (2012). A hybrid Data Mining Technique for Improving the Classification accuracy of Microarray Dataset. *I.J. Information engineering and electronics Business*, 2, 43-50.

DeBock, K. W., Coussement, K., & VandenPoel, D. (2010). Ensemble classification based on generalized additive models. *Computational Statistics & Data Analysis*, *54*(6), 1535–1546. doi:10.1016/j.csda.2009.12.013

Dietterich, T. (2000). Ensemble methods in machine learning. *Proceedings of the Multiple Classifier System Conference* (pp. 1– 15). 10.1007/3-540-45014-9_1

Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, *40*(2), 139–157. doi:10.1023/A:1007607513941

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. doi:10.1006/jcss.1997.1504

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*, 1157–1182.

Hall, M. A. (2000). Correlation-based feature selection for discrete and numeric class machine learning. *Proc. Seventeenth International Conference on Machine Learning* (pp. 359–366).

Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(10), 993–1001. doi:10.1109/34.58871

Hsu, H.-H., Hsieh, C.-W., & Lu, M.-D. (2011). Hybrid feature selection by combining filters and wrappers. *Expert Systems with Applications*, *38*(7), 8144–8150. doi:10.1016/j.eswa.2010.12.156

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., ... Meltzer, P. S. (2001). Classification and diagnostic of cancers using gene expression profiling and Artificial neural networks. *NCBI*, *7*(6), 673–679. PMID:11385503

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1-2), 273–324. doi:10.1016/S0004-3702(97)00043-X

Kuncheva, L. I., & Rodriguez, J. J. (2007). An experimental study on Rotation Forest ensembles. Proceedings of *MCS 2007, LNCS* (Vol. 4472, pp. 459–468). Berlin: Springer. doi:10.1007/978-3-540-72523-7_46

Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles. *Machine Learning*, *51*(2), 181–207. doi:10.1023/A:1022859003006

Liu, H., Liu, L., & Zhang, H. (2010). Ensemble gene selection for cancer classification. *Pattern Recognition*, *43*(8), 2763–2772. doi:10.1016/j.patcog.2010.02.008

Liu, H., Motoda, H., & Yu, L. (2002). Feature selection with selective sampling. *Proc. Nineteenth International Conference on Machine Learning* (pp. 395–402).

Liu, K.-H., & Huang, D.-S. (2008). Cancer classification using Rotation Forest. *Computers in Biology and Medicine*, *38*(5), 601–610. doi:10.1016/j.compbiomed.2008.02.007 PMID:18394595

Osareh, A., & Shadgar, B. (2013). An efficient ensemble learning method for gene microarray classification. Biomed Research International, 2013.

Peng, Y. (2006). A novel ensemble machine learning for robust microarray data classification. *Computers in Biology and Medicine*, *6*(36), 553–573. doi:10.1016/j.compbiomed.2005.04.001 PMID:15978569

Petricoin, E. F. III, Ardekani, A. M., Hitt, B. A., Levine, P. J., Fusaro, V. A., Steinberg, S. M., ... Liotta, L. A. (2002). Use of Proteomic Patterns in Serum to Identify Ovarian Cancer. *Lancet*, *359*(9306), 572–577. doi:10.1016/S0140-6736(02)07746-2 PMID:11867112

Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., ... Golub, T. R. (2002). Prediction of central nervous system embryonal tumor outcome based on gene expression. *Nature*, *415*(6870), 436–442. doi:10.1038/415436a PMID:11807556

Thomas, A., & Thibault, H. (2010). Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics (Oxford, England)*, *26*(3), 392–398. doi:10.1093/bioinformatics/btp630 PMID:19942583

Yang, K., Cai, Z., Li, J., & Lin, G. (2006). A stable gene selection in microarray data analysis. *BMC Bioinformatics*, *7*(1), 228. doi:10.1186/1471-2105-7-228 PMID:16643657

Yang, P., Yang, Y. H., Zhou, B. B., & Zomaya, A. Y. (2010). A review of ensemble methods in bioinformatics. *Current Bioinformatics*, *5*(4), 296–308. doi:10.2174/157489310794072508

Yongjun, P., Minghao, P., Kiejung, P., & Keun, H. R. (2012). An ensemble correlation-based gene selection algorithm for cancer classification with gene expression data, *Data and Text Mining*. *Bioinformatics (Oxford, England)*, *28*(24), 3306–3315. doi:10.1093/bioinformatics/bts602 PMID:23060613

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yu, L., & Liu, H. (2003). Feature selection for high-dimensional data: a fast correlation- based filter solution, Proc. *Proceedings of the Twentieth International Conference on Machine Learning* (pp. 856–863).

Zhang, C.-X., & Zhang, J.-S. (2008). RotBoost: A technique for combining rotation forest and Adaboost. *Pattern Recognition Letters*, *29*(10), 1524–1536. doi:10.1016/j.patrec.2008.03.006

Zhang, C.-X., & Zhang, J.-S. (2010). A variant of rotation forest for constructing ensemble classifiers. *Pattern Analysis & Applications*, *13*(1), 59–77. doi:10.100710044-009-0168-8

KEY TERMS AND DEFINITIONS

Base Learners: The component / individual learner of the ensemble which are combined together strategically is referred to as base learners. Base (weak) learner must focus on correctly classifying the most highly weighted examples while strongly avoiding over-fitting.

Correlation-Based: Fluctuation of one variable reliably predicts a similar fluctuation in another variable i.e., change in one causes the change in other.

Diversity: The relation that holds between two entities / objects when and only when they are not identical; the property of being numerically distinct.

Ensemble Learning: It is the process by which multiple models, such as classifiers or experts are strategically generated and combined using a particular rule to solve complex computational intelligence problem.

Feature Redundancy (FR): Duplicated features or information, that adds as a precaution against failure or error.

Hybrid Ensemble: Combining two different ensemble models to enhance the prediction/ generalization capability of the ensemble model.

Principal Component Analysis: It's a method of analysis which involves finding the linear combination of a set of variables that has maximum variance and removing its effect and repeating this successively.

Random Projection: In mathematics and statistics, random projection is a technique used to reduce the dimensionality of a set of points which lie in Euclidean space.

This research was previously published in the Handbook of Research on Computational Intelligence Applications in Bioinformatics edited by Sujata Dash and Bidyadhar Subudhi; pages 17-36; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 39
Fall Detection Depth-Based Using Tilt Angle and Shape Deformation

Fairouz Merrouche
University of Science and Technology USTHB, Algeria

Nadia Baha
University of Science and Technology USTHB, Algeria

ABSTRACT

The population of elderly people is in growth. Falls risk their life, to disabilities, and to fears. Automatic fall detection systems provide them secure living; helping them to be independent at home. Computer vision offers efficient systems over many developed systems. In this article, the authors propose a new vision-based fall detection using depth camera. It combines human shape analysis, centroid detection and motion where it exploits the 3D information provided by a Kinect to compute the tilt angle to discriminate falls. Experimental tests were done with SDUFall dataset that contains 20 subjects performing five daily activities and falls, demonstrate the efficiency of the proposed system, and show that our method is promising achieving satisfactory results up to 84.66%.

INTRODUCTION

In recent years, the increase of life expectancy has led to the growth of aging population. Independent living and not needing help can lead to severe consequences. Falling are risky event that can occur in an elder person life and repeated falling are the major danger ("Health Organization Department," 2008), hospitalization and traumatic injury-related deaths (Stevens and Rudd, 2010) for person aged 65 or elder. Furthermore, fall can have physiological consequences such depress, avoidance of activities, a fear of falling that increase from suffering (Friedman et al., 2002).

DOI: 10.4018/978-1-7998-1204-3.ch039

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

There have been various researches, in order to overcome elderly fall risks and prevent these incidents. Some of these technologies are based on wearable sensors such as accelerometers, gyroscope (Cheng et al., 2013). Nevertheless, these sensors are ineffective when the user forgets to wear them or forgets to recharge or to replace the battery.

Vision-based technologies that use commodity cameras have proven to offer better solutions. Cameras have a low price, and not required to be worn by the user (Rougier et al., 2011).

Recently, Kinecth has received much attention of lot of researches due to its low-cost price and many other advantages against regular RGB camera like the privacy of person's life. In this paper, we will present a new vision-based method for fall detection in an indoor environment. The main contribution of this work lies on detecting falls differently in simple way by using the tilt angle based on head tracking and centroid movement.

In this paper, we outline a new vision-based approach using depth camera Kinect to detect falls and exploiting person shape based on tilt angle and motion.

The remainder of this paper is organized as follow section 2 presents the related work in the field of fall detection systems. Section 3 presents the stages followed to fall detection. In section 4, experimental results are presented and discussed. Finally, section 5 concludes the paper with some remarks.

RELATED WORK

In the last few years, various approaches have been put forward by many researches to improve fall detection solutions. Fall detection methods are subdivided into two classes: computer vision-based methods and non-computer vision method.

In this section, we will review some works proposed for fall detection starting by non-vision category. Many sensors have been exploited. In the work of (Cheng et al., 2013) the authors have attached two accelerometers to chest and thigh of 10 persons. Decision tree was applied to recognize posture transitions with threshold to detect falls. The main disadvantage of this method is to wear the sensor.

Another work was proposed in (Sannino et al. 2015) where a wearable accelerometer and a mobile device were used to monitor the patients in real time, the data recorded by the accelerometer is sent to the mobile device to analyze them. This approach has three phases. First, gather data and make a set of simulated falls with the corresponding class "fall" or "non-fall". Second, knowledge extraction as a set of IF-THEN rules and the last phase consists in patients monitoring in the decisional layer in real time but this approach still preliminary, the tests were done for just three people.

In the work of (Alwan et al., 2006), a special piezoelectric sensor coupled to the floor is used, a binary fall signal can be generated in case of fall, and however, the main drawbacks can be false alarms.

In (Yazar et al., 2014) fall was detected based on three sensors consisting of a vibration sensor and two PIR sensors. The vibrations were converted into electrical signals and then Feature vectors from the vibration waveforms are extracted using the complex wavelet transform (CWT) and they are classified using support vector machine (SVM). PIR sensors were used as additional sensors to detect the infrared radiation emitted by moving objects in the room and that to eliminate falls alarms generated when the vibration sensor signal energy is concentrated in low-frequency bands such as vibration from door slams and other similar events.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In recent work in (Jokanovic et al., 2016) a new method was proposed based on radar technologies to detect falls, the authors proposed to use time-frequency (TF) which is able to reveal higher order velocity for different part of human body and based on deep learning which learns and captures the properties of the TF signatures without human intervention then it is fed to a classifier achieving 87% success rate.

These solutions are promising but cannot be effective due to noise they generate and cannot satisfy the elder user due to the inconvenience they provide why the computer vision-based solution has been adopted by many researches. There are some methods that used shape analysis to detect fall. Generally, the first step is to detect the person in each video frame using separation of people from the background. In 2D sensors, the bounding box aspect ratio was used to distinguish if the person is in an upright position or not (Liu et al., 2010).

In (Foroughi et al., 2008), an ellipse approximated the shape of the person. Despite their performance, these approaches are not feasible for fall into the direction of the camera, therefore a multi camera system was proposed to overcome this problem which benefits from the 3D features to detect fall.

(Worrakulpanit and samanpiboon, 2014) also proposed a fall detection method based the orientation standard deviation of the Ellipse, so the silhouette s approximated by an ellipse and then the motion is quantified by calculating pixel value of Motion History Image. Fall is discriminated from other activities by using standard deviation of C-Motion. According to the author this method is robust.

In (Anderson et al., 2008) multi-camera system is used where fall detection has been presented in two levels. The first level infers the states of the object in each frame. The second level consists in linguistic summarization of 3D person's states called voxels.

(Rougier et al., 2011) proposed another approach where they used four cameras were mounted in one room for classification using GMM to detect fall by analyzing the deformation in human shape. (A. Mousse et al. 2016) propose a multi-camera system based on two relatively orthogonal views to measure the percentage of the orthogonal views of the person which is in contact with ground according to the information provided by each camera and this to differentiate between lying down activity and other activities the results were promising 95.8% sensitivity and 100% specificity.

Recently, the 3D camera has attracted a lot of attention and the use of Kinect is widely emerged because of its low cost. In (Stone and Skubic, 2015) the centroid height of the person relative to the ground and centroid velocity were used to detect occluded fall. The person was tracked over time to obtain persons vertical state to segment on ground events and then an ensemble of decision trees to compute a confidence that a fall occurs before an on-ground state.

In (Bian et al., 2015) the key joint of human body is extracted using an RDT algorithm trained with Shannon entropy and then track these joints. SVM classier is employed to detect fall based on head joint distance trajectory. A new shape-based method to detect fall is used in (Ma et al, 2014) where CSS features of the silhouette in each frame and the actions are represented by a bag of CSS words and then fall is detected with classification using VPSO-ELM that uses particle swarm optimization.

In (Kwolek and Kepski, 2015) a new fall detection method is presented combining an accelerometer with Kinect. The accelerometer was used to filter out a non-fall events and Kinect sensor to authenticate the fall, which leads to good performance and another work of (Pramerdorfer et al. 2016) where they emphasize person detection and tracking using features that invariant with respect to sensor position, and detect falls in three steps state prediction, event detection and fall verification. They tested their method on experimental condition and they had good results and robust to partial occlusions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(Wang et al., 2016) proposed an automatic fall detection method where they combine Histograms of Oriented Gradients (HOG), Local Binary Pattern (LBP) features extracted by deep learning Framework Caffee to form a new feature HLC. They achieve an average result of 93.7% sensitivity and 92.0% specificity on three datasets.

In the work of (Baha et al., 2016) another work was proposed using two Kinect sensors. The first one is frontal and the other is on ceiling. The first one is based on the person's depth information and his velocity by computing the surface of the bounding box and then measure surface difference and velocity to detect fall and validate it. The second one is based on the variation of bounding box parameters and its velocity (frontal Kinect) and fall is detected based on the value of width to height of the bounding box and also to velocity. A combination of the two methods was performed to improve accuracy achieving up 95%.

In (Fan et al., 2017) a normalized directional histogram is developed around the center of the ellipse to represent a human posture by multidirectional statistical analysis and then 12 static and 8 dynamic features where extracted these features were fed into a directed acyclic graph support vector machine to distinguish four human postures A fall-like accident is detected by counting the occurrences of lying postures in a short temporal window and a fall event is detected by immobility verification.

In (Li et al., 2018) a combination of Kinect and mobile phone is exploited. Two methods are proposed which are based on threshold for mobile phone and support vector machine for Kinect. Fall is detected using the fusion of the data provided from the two subsystems based on two methods the logical rules-based and D-S evidence fusion this approach achieves 100% of accuracy.

In (Cameron et al. 2018) the authors propose fall detection based on machine learning method. They used gradient-based local visual feature extraction including histogram of oriented gradients (HOG), histogram of motion gradients (HMG), histogram of optical flow (HOF), and motion boundary histograms (MBH). They provide a comparative study with different scenario and they demonstrated the effectiveness of this approach.

In this work, we propose a novel human shape-based fall detection that makes use of Kinect and benefits from its advantage against RGB camera and we use shape information based on the angle of the line linking the head and the center of mass of the person.

PROPOSED FALL DETECTION METHOD

The proposed system uses Kinect sensor which is a special camera developed by Microsoft for Xbox360 console, first introduced in 2010. This sensor is composed of an RGB camera and an infrared (IR) depth sensor, which are characterized by a resolution of 640x480 at 30 fps.

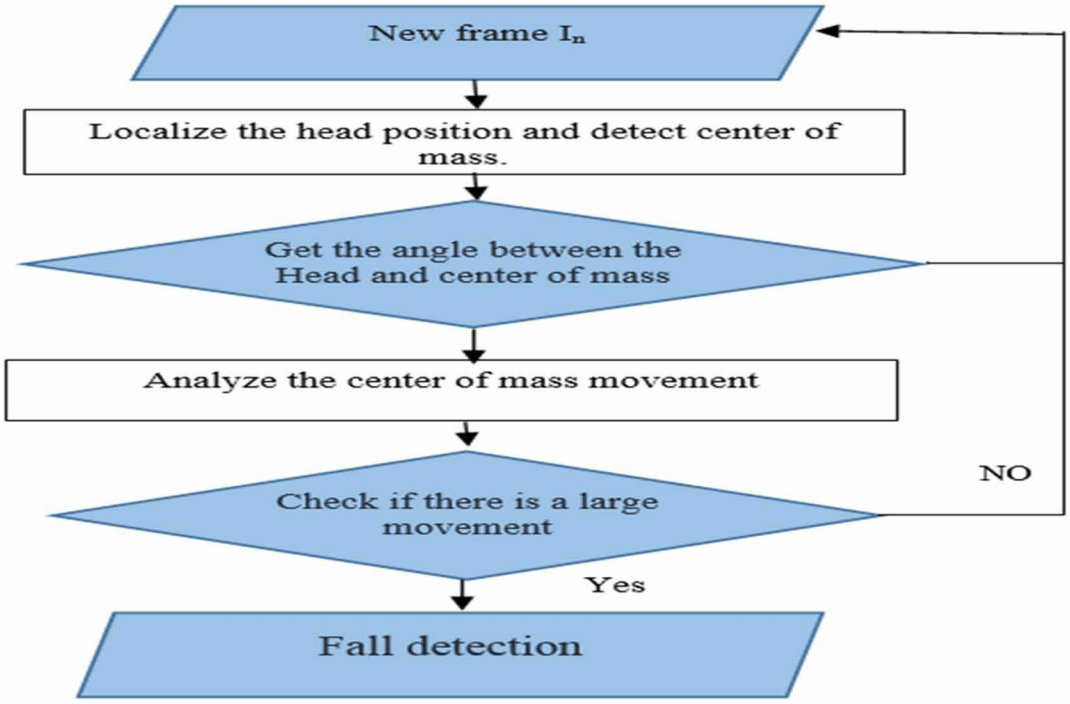
The Kinect is able to provide a depth map where each pixels gray scale represents the distance in millimeters between the Kinect and the objects. The infrared sensor is able to work even in a dark room, which represents an advantage for any surveillance system. In this work, we will exploit the advantage of this sensor and use only the depth map to respect the privacy of the elder people. Figure 1 shows an image of Microsoft Kinect.

In the following sections, we will describe in more details how we use human shape analysis and movement for fall detection. In the first time, we start with preprocessing stage. Second, we localize the position of the head at each frame using a tracking method, then, we measure the position of the centroid, after that we compute the tilt angle. Finally, fall is detected and validated. Figure 2 shows an overview of fall detection method.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Microsoft Kinect

Figure 2. Overview of fall detection method



Usually, most computer vision systems need preprocessing steps to have accurate fall detection. In the following subsections, we will detail these steps.

Background Subtraction

Getting the human body silhouette is an important step in our system. A Gaussian mixture model GMM background subtraction method is used to extract the foreground where each pixel is characterized by a mixture of k Gaussian. Every pixel is evaluated for its probability whether it is belonging to Background or to Foreground, then morphological operators: erosion and dilatation are applied to discover any left pixels belonging to the foreground or to the background. Figure 3 shows an example of an image before and after background subtraction step.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Background subtraction example

Head Tracking

Once the silhouette obtained, a tracking method is used in order to localize the head at each frame. In this work, the head is approximated by an ellipse and tracked using particle filter. Particle Filter is widely used for tracking because of its performance and its complexity. It employs a Sequential Importance Sampling algorithm or Monte Carlo method to compute the posterior density of a variable given an observation (Gordon et al. 1993), the posterior density function PDF over the state can be approximated by a set of particles. Each particle x_i^t represents a state with associated weights w_i^t which is the observation likelihood:

$$w = p(z^t \mid w^t) \tag{1}$$

where z^t is an observation measurement.

In our case, the observation combines the score between the ellipses and head point. The tracking algorithm has three steps, first, the initialization step; second the update and then the resampling step. Algorithm 1 in Figure 4 illustrates the use of particle filter for head tracking and an example of use of tracker is illustrated in Figure 5.

Tilt Angle

Some works in the literature used the tilt angle to detect falls (Foroughi et al. 2008). But they use generally the 2D information where the angle is computed in the x-y plan. Contrary, in the proposed method, we will present another view of the angle where the 3D information provided by the kinect sensor will be exploited to compute the angle. We agree that the y direction gives us the information about the vertical deformation of the human shape and the z direction (axis) gives us the other very important information deformation along depth. Therefore the tilt angle is computed as the angle between the projection of the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure4. Head tracking algorithm

Input: Ellipses m odeling the head.

Output: Position of the head.

begin

n← 0; i← 0

for i=0 to m do:

- Spread m ellipses with initial weights w_t^i

end

for i=0 to m do:

- Update the weights by the likelihood:

$$w_t^i = w_{t-1}^i \; p(y_t^i \mid x_t^i)$$

Norm alize to:

$$w_t^i = w_t^i / \sum w_t$$

end

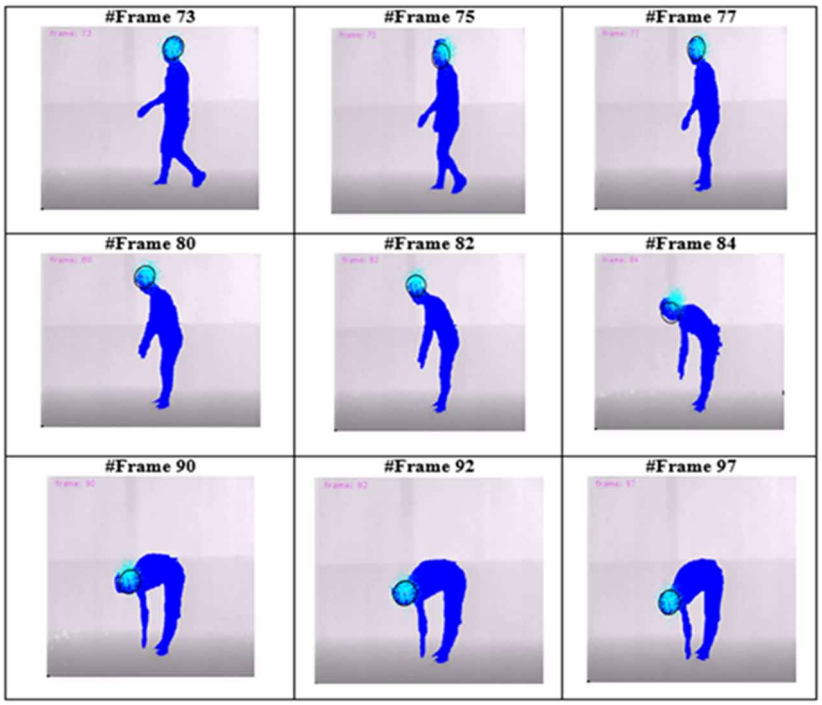
for i=0 to m do:

- Choose the reasonable particles based on the likelihood of each of them.

end

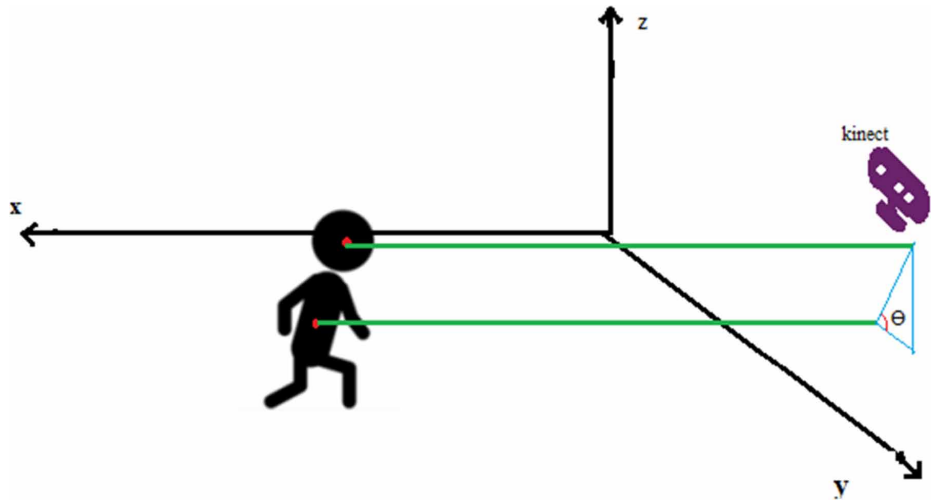
end

Figure 5. Head tracking in depth frames



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The angle θ between the head and centroid



head point and centroid point P_h (x_h, y_h, z_h) and P_c (x_c, y_c, z_c) respectively on the y-z plan as illustrates in Figure 6.

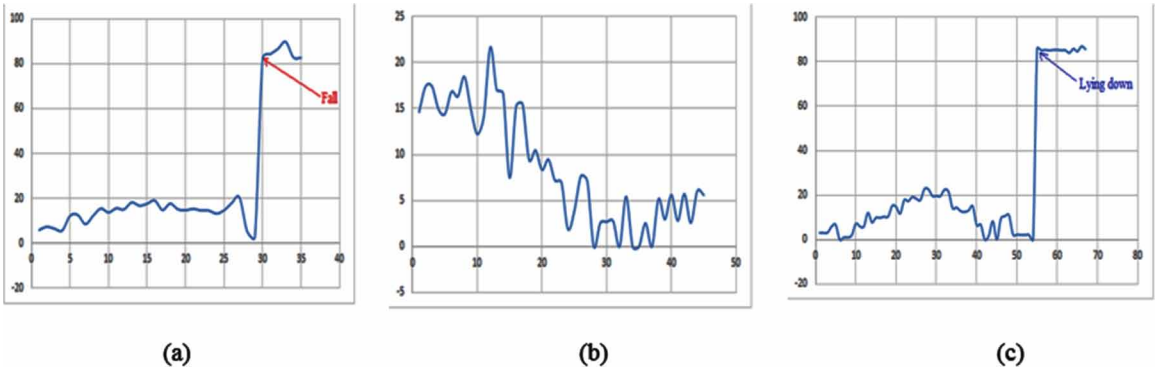
The tilt angle θ is computed using the following formula:

$$\theta = arctan \frac{z_h - z_c}{y_h - y_c} \tag{2}$$

When the angle θ exceeds a certain threshold Thf, lying down pose and squatting pose event which are very similar to lying down pose are detected.

In Figure 7, we represent the angle variation, which characterizes a fall, lying down and squatting events.

Figure 7. Angle variation in the case of (a) Fall event, (b) Squatting, (c) Lying down event



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

To distinguish between the three events (falling, lying down or squatting), motion information will be exploited. As the motion of the centroid changes smoothly during human movements as well as when there is a sudden and rapid movement like fall. The displacement of the centroid movement will be measured. Therefore, we will adopt the method of (Merrouche and Baha, 2017) and we measure the position of the centroid at each frame.

We generate a curve t_c tracing the centroid movement at the beginning of a suspicious event (falling, squatting, lying) by going backward t seconds. The t value represents the time average that falling event can take to occur, less than one second (Rougier et al., 2007), in our implementation, we fixed it to ~ 0.83 seconds.

We will compare the resultant curve to the free fall curve. We generate this curve using the next formula:

$$z_t = z_0 - 0.5a\left(t - t_0\right)^2 \tag{3}$$

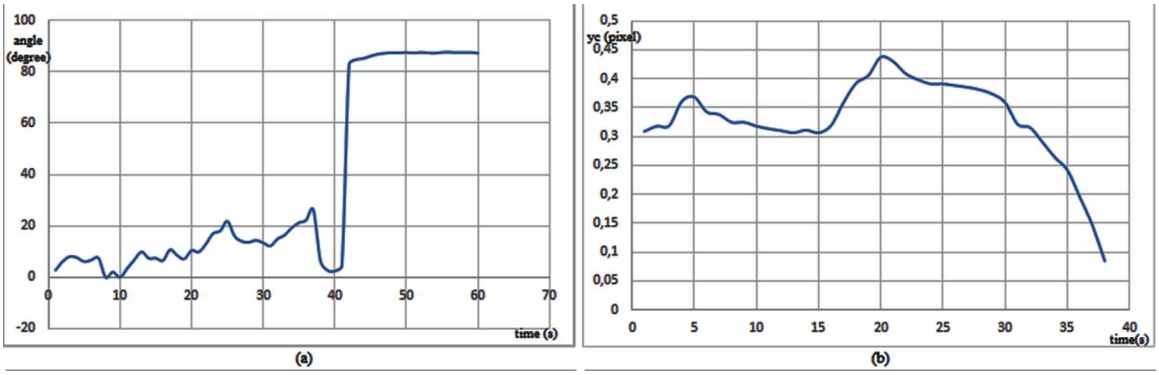
where z_t represents the centroid ordinate at each time t and z_0 represents the centroid ordinate at t_0. t_0 defines the beginning of the suspicious event. a is the acceleration.

Generally, when an older person squats, he squats slowly but in the data set (Ma et al. 2014), that we use, some actors simulates squatting action with large movement, to overcome these false alarms it is necessary to adopt another feature. For that, we choose the distance to the floor due to its discriminant propriety. Figure 8a represents angle variation for squatting event that is returned as fall alarm and Figure 8b represents its velocity.

Figure 9 illustrates an example of fall detection obtained for person 9 of SDUFall dataset (Ma et al., 2014) falling down.

The steps detailed previously are summarized in algorithm 2 in Figure 10.

Figure 8. (a) Angle variation for squatting event, (b) Its velocity



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. An example of fall detection

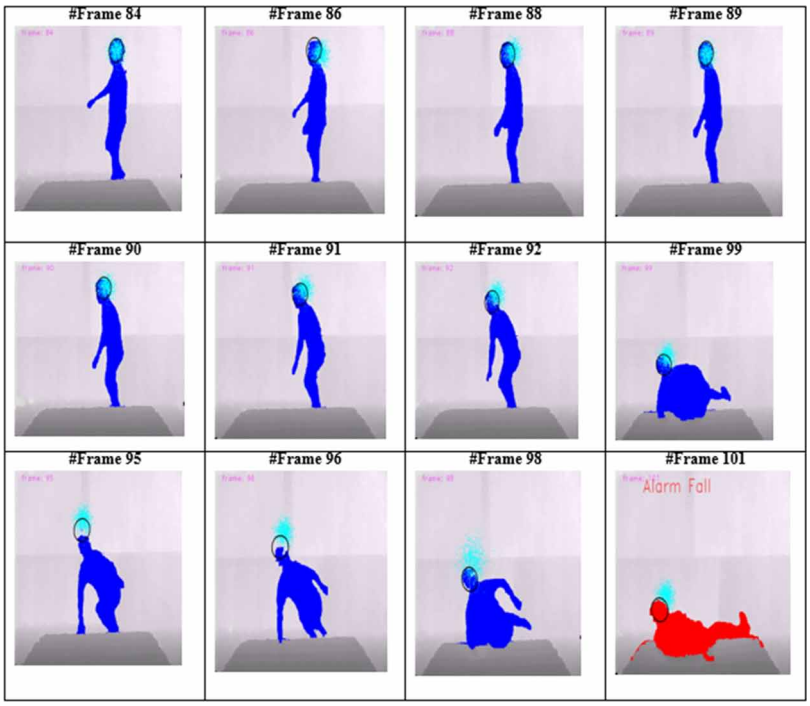


Figure10. Fall detection process algorithm

Input: Frames sequence

Output: Fall or non-fall.

Begin

1) Background subtraction.
2) Head Tracking using particle filter.
3) Compute the angle θ between the centroid and the Head.
4) If $\theta > thr1$ then
   If there is a large movement then
      If distance to the floor $> thr2$ then
   Fall validated.
      End If
   End If
End If

End.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

EXPERIMENTAL RESULTS AND DISCUSSION

In this section, different experiments conducted are detailed.

Data Base

To test the performance of our method, we use SDUFall dataset available at http://www.sucro.org/homepage/wanghaibo/SDUFall.html. The SDUFall Dataset contains several actions of 20 person men and women. These data were recorded with Kinect installed at 1.5 heights where sitting down, lying down, bending, squatting and falling actions were performed 10 times with different condition like carrying or not an object, turning the light on and off and changing the direction towards the camera. These conditions were chosen randomly.

In total, there are 1200 depth video with frame size of 320x240 recorded at 30 fps in avi format.

Results and Discussion

All experiments were conducted on a PC with Intel (R) Core i5 1.80 GHz. Algorithms were implemented using C++ language and openCV library. We used the following parameters to evaluate the proposed method of fall detection (Bian et al., 2015):

TP: True positives are the number of fall events detected as fall.
TN: True negatives are the number of non-fall events detected as non-fall events.
FP: False positives are the number of non-fall events detected as fall events.
FN: False negatives are the number of fall events detected as non-falls events.
Sensitivity: Indicates the capacity of the algorithm to detect fall events computed as:

$$sensitivity = \frac{TP}{TP + FN} \tag{4}$$

Specificity: indicates the capacity of the algorithm to detect non-fall events computed with following formula:

$$specificity = \frac{TP}{TN + FP} \tag{5}$$

Accuracy: Indicates how many true results were correctly detected amongst all measurements computed as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The results show the performance of the proposed method to detect fall alarms and its ability to discriminate between falls, lying down events and squatting. Four people from SDU-Fall dataset were chosen here for demonstration. In Table 1, 2, 3 and 4, we present the three actions falling, lying down and squatting performed 10 times. We can observe from Table 1, 2, 3 and 4, that the result for true positive is 100%. The specificity and sensitivity varies between 70% and 100%.

Although the proposed method presents good results, however, in some cases, the system presents some errors. There were some false negatives, the most likely explanation of these results is the errors that occur in head tracking step and some other false positives occurred when a person brutally lies down. Other error detections were due to errors in background subtraction step for that, in the future we must use more robust background subtraction method. Figure 11(a) shows an example of false alarm due to error with background subtraction and Figure 11 (b) shows an example how the tracker missed the head.

Table 1. Experimental results with person 3

	Falling	**Squatting**	**Lying Down**
FN	0		
FP		2	1
TP	10		
TN		8	9
Specificity		80%	90%
Sensitivity	100%		

Table 2. Experimental results with person 14

	Falling	**Squatting**	**Lying Down**
FN	2		
FP		1	1
TP	8		
TN		9	9
Specificity		90%	90%
Sensitivity	80%		

Table 3. Experimental results with person 18

	Falling	**Squatting**	**Lying Down**
FN	0		
FP		1	3
TP	10		
TN		9	7
Specificity		90%	70%
Sensitivity	100%		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Experimental results with person 11

	Falling	Squatting	Lying Down
FN	0		
FP		2	1
TP	10		
TN		8	9
Specificity		80%	90%
Sensitivity	100%		

Figure 11. (a) False alarm due to error in the returned background subtraction. (b) False negative due to missing the head.

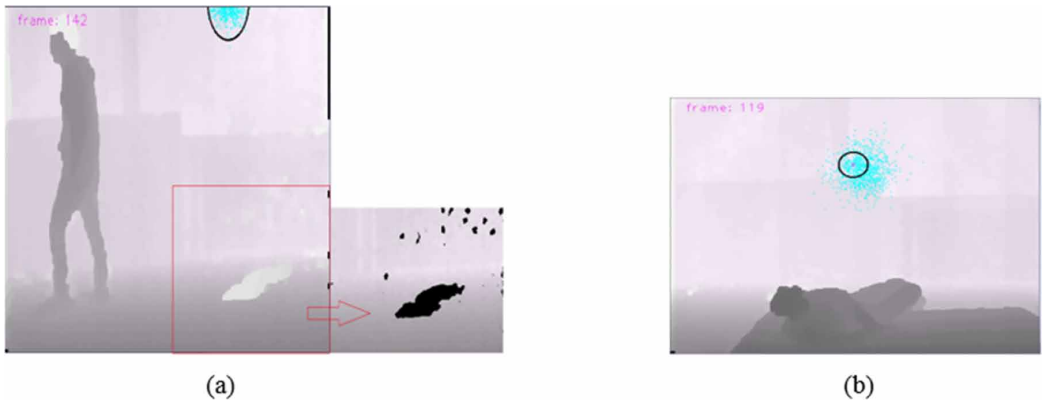


Figure 12 represents the results obtained with all the events in SDUFall DataSet and Table 5 shows the performance obtained by the proposed approach.

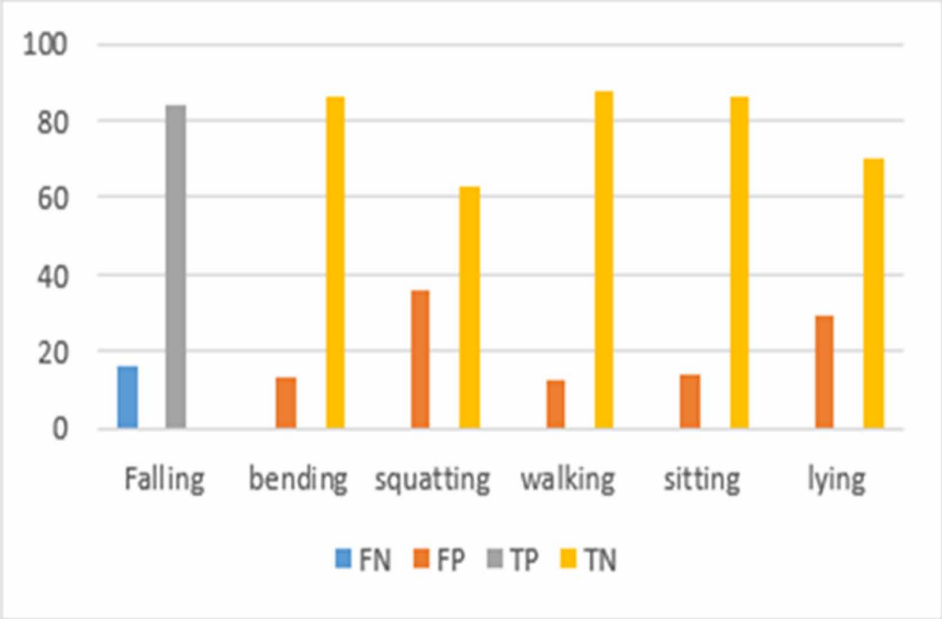
In this paper, the focus attention was to demonstrate the robustness of the use of the tilt angle to discriminate falls. Compared to the state of the art, we computed the angle in a simplified way. Also we exploited the data recorded by Kinect to detect backward an afterward falls as the angle was measured on the y-z plan unlike the work of (Rougier et al., 2007) where the angle represents the angle of a fitting ellipse in x-y plan and in contrast to the work of (Yang et al., 2014) where the angle was represented as the angle between the central line and y axis. Moreover, the results show that the proposed system is able not only to detect falls, but also to differentiate between falls, lying down and squats efficiently, and without forgetting the advantage of Kinect to protect the elder person privacy.

Table 5. Fall detection performance for SDUFall dataSet

Accuracy (%)	Specificity (%)	Sensitivity (%)
84.66	84.8	84

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Results with different events SDUFalldataset



CONCLUSION

In this paper, we have presented a novel fall detection method using Kinect sensor. The method combines shape features and centroid movement. The head has been tracked using particle filter.

Centroid was detected at each frame. The angle between the head and the centroid was measured. The movement of the centroid was analyzed which allowed us to distinguish between lying down, squatting and fall events.

In this paper, we demonstrated the robustness of our method. The conducted experiments showed that our method is efficient and can detect falls accurately compared to the state of the art. The accuracy has been achieved up to 84.66%. Further work needs to be performed to compute the tilt angle on both x-y plane and y-z plane and combine them to improve accuracy.

As our results are encouraging should be validated by a larger sample size. We will address to have complicated behaviors with more scenarios such as backwards fall and fall from chair.

REFERENCES

Alwan, M., Rajendran, P. J., Kell, S., Mack, D., Dalal, S., Wolfe, M., & Felder, R. (2006, April). A smart and passive floor-vibration based fall detector for elderly. In 2nd Information and Communication Technologies ICTTA'06 (Vol. 1, pp. 1003-1007). IEEE. doi:10.1109/ICTTA.2006.1684511

Anderson, D., Luke, R. H., Keller, J. M., Skubic, M., Rantz, M., & Aud, M. (2009). Linguistic summarization of video for fall detection using voxel person and fuzzy logic. *Computer Vision and Image Understanding*, *113*(1), 80–89. doi:10.1016/j.cviu.2008.07.006 PMID:20046216

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Baha, N., Beloudah, E., & Ousmer, M. (2016). Towards a real-time fall detection system using Kinect Sensor. *International Journal of Computer Vision and Image Processing*, *6*(1), 41–58. doi:10.4018/IJCVIP.2016010104

Bian, Z. P., Hou, J., Chau, L. P., & Magnenat-Thalmann, N. (2015). Fall detection based on body part tracking using a depth camera. *IEEE Journal of Biomedical and Health Informatics*, *19*(2), 430–439. doi:10.1109/JBHI.2014.2319372 PMID:24771601

Cameron, R., Zuo, Z., Sexton, G., & Yang, L. (2017, September). a fall detection/recognition system and an empirical study of gradient-based feature extraction approaches. In Inuk workshop on computational intelligence (pp. 276-289). Cham: Springer.

Cheng, J., Chen, X., & Shen, M. (2013). A framework for daily activity monitoring and fall detection based on surface electromyography and accelerometer signals. *IEEE Journal of Biomedical and Health Informatics*, *17*(1), 38–45. doi:10.1109/TITB.2012.2226905 PMID:24234563

Fan, K., Wang, P., Hu, Y., & Dou, B. (2017). Fall detection via human posture representation and support vector machine. *International Journal of Distributed Sensor Networks*, *13*(5), 1550147717707418. doi:10.1177/1550147717707418

Foroughi, H., Rezvanian, A., & Paziraee, A. (2008, December). Robust fall detection using human shape and multi-class support vector machine. In *Sixth Indian Conference on Computer Vision, Graphics & Image Processing ICVGIP'08* (pp. 413-420). IEEE. 10.1109/ICVGIP.2008.49

Friedman, S. M., Munoz, B., West, S. K., Rubin, G. S., & Fried, L. P. (2002). Falls and fear of falling: Which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. *Journal of the American Geriatrics Society*, *50*(8), 1329–1335. doi:10.1046/j.1532-5415.2002.50352.x PMID:12164987

Gordon, N. J., Salmond, D. J., & Smith, A. F. (1993, April). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings. Part F. Radar and Signal Processing*, *140*(2), 107–113. doi:10.1049/ip-f-2.1993.0015

Health Organization Department of Ageing and Life Course. (2008). *WHO global report on falls prevention in older age*. Geneva, Switzerland: World Health Organization.

Jokanovic, B., Amin, M., & Ahmad, F. (2016, May). Radar fall motion detection using deep learning. In *2016 IEEE Radar Conference (RadarConf)* (pp. 1-6). IEEE. 10.1109/RADAR.2016.7485147

Kwolek, B., & Kepski, M. (2015). Improving fall detection by the use of depth sensor and accelerometer. *Neurocomputing*, *168*, 637–645. doi:10.1016/j.neucom.2015.05.061

Li, X., Nie, L., Xu, H., & Wang, X. (2018). collaborative fall detection using smart phone and Kinect. *Mobile Networks and Applications*, 1–14.

Liu, C. L., Lee, C. H., & Lin, P. M. (2010). A fall detection system using k-nearest neighbor classifier. *Expert Systems with Applications*, *37*(10), 7174–7181. doi:10.1016/j.eswa.2010.04.014

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ma, X., Wang, H., Xue, B., Zhou, M., Ji, B., & Li, Y. (2014). Depth-based human fall detection via shape features and improved extreme learning machine. *IEEE Journal of Biomedical and Health Informatics*, *18*(6), 1915–1922. doi:10.1109/JBHI.2014.2304357 PMID:25375688

Merrouche, F., & Baha, N. (2017, July). Fall Detection using Head Tracking and Centroid Movement Based on a Depth Camera. In *Proceedings of the International Conference on Computing for Engineering and Sciences* (pp. 29-34). ACM. 10.1145/3129186.3129192

Mousse, M. A., Motamed, C., & Ezin, E. C. (2017). Percentage of human-occupied areas for fall detection from two views. *The Visual Computer*, *33*(12), 1529–1540. doi:10.100700371-016-1296-y

Pramerdorfer, C., Planinc, R., Van Loock, M., Fankhauser, D., Kampel, M., & Brandstötter, M. (2016, October). Fall Detection Based on Depth-Data in Practice. In *European Conference on Computer Vision* (pp. 195-208). Cham: Springer. 10.1007/978-3-319-48881-3_14

Rougier, C., Auvinet, E., Rousseau, J., Mignotte, M., & Meunier, J. (2011, June). Fall detection from depth map video sequences. In *International Conference on Smart Homes and Health Telematics* (pp. 121-128). Springer, Berlin, Heidelberg. 10.1007/978-3-642-21535-3_16

Rougier, C., Meunier, J., St-Arnaud, A., & Rousseau, J. (2007, May). Fall detection from human shape and motion history using video surveillance. In *21st International Conference on Advanced Information Networking and Applications Workshops AINAW'07* (Vol. 2, pp. 875-880). IEEE. 10.1109/AINAW.2007.181

Rougier, C., Meunier, J., St-Arnaud, A., & Rousseau, J. (2011). Robust video surveillance for fall detection based on human shape deformation. *IEEE Transactions on Circuits and Systems for Video Technology*, *21*(5), 611–622. doi:10.1109/TCSVT.2011.2129370

Sannino, G., De Falco, I., & De Pietro, G. (2015). A supervised approach to automatically extract a set of rules to support fall detection in anmHealth system. *Applied Soft Computing*, *34*, 205–216. doi:10.1016/j.asoc.2015.04.060

Stevens, J. A., & Rudd, R. A. (2014). Circumstances and contributing causes of fall deaths among persons aged 65 and older: United States, 2010. *Journal of the American Geriatrics Society*, *62*(3), 470–475. doi:10.1111/jgs.12702 PMID:24617970

Stone, E. E., & Skubic, M. (2015). Fall detection in homes of older adults using the Microsoft Kinect. *IEEE Journal of Biomedical and Health Informatics*, *19*(1), 290–301. doi:10.1109/JBHI.2014.2312180 PMID:24733032

Wang, K., Cao, G., Meng, D., Chen, W., & Cao, W. (2016, December). Automatic fall detection of human in video using combination of features. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1228-1233). IEEE.

Worrakulpanit, N., & Samanpiboon, P. (2014). Human fall detection using standard deviation of C-motion method. *Journal of Automation and Control Engineering*, *2*(4), 388–391. doi:10.12720/joace.2.4.388-391

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yang, S. W., & Lin, S. K. (2014). fall detection for multiple pedestrians using depth image processing technique. *Computer Methods and Programs in Biomedicine*, *114*(2), 172–182.

Yazar, A., Erden, F., & Cetin, A. E. (2014, May). Multi-sensor ambient assisted living system for fall detection. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'14)* (pp. 1-3).

This research was previously published in the International Journal of Computer Vision and Image Processing (IJCVIP), 8(4); edited by Jose Garcia-Rodriguez; pages 26-40; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 40
Energy Efficient Particle Optimized Compressed ECG Data over Zigbee Environment

Dilip Kumar
SLIET, India

Rajeev Kumar
DAVIET, India

Tony Singla
DAVIET, India

ABSTRACT

Standard Electrocardiogram Tracking instruments are huge to carry away over the remote areas for the surveillance. Holter is the compact instrument meant for collecting ECG of a patient without a pause while the patient is on the go of their daily activities. Holter works on battery for 48 hours without any angle of transmission but when allowed to transmit battery power dies soon, for these purposes some energy saving techniques is required. In this chapter the authors have proposed a Wavelet based Compression Technique, followed by Optimization under Genetic Algorithm and Particle Swarm Optimization. Compressed and Optimized ECG data has been transferred over Zigbee IEEE 802.15.4 with the intention of saving energy implicating it on a hardware chip. Transferred data will be available to the Doctor for on time treatment and further examination and storage. Embedded prior techniques in Holter can enhance its life, with fact of sending crucial data.

INTRODUCTION

Present time reveals the truth that there is a raise in the aging population, busy world and life taking diseases (Maneesha V. Ramesh et al., 2012). Ongoing era demands the early detection and prevention of death threatening diseases [2], out of a huge number here the main focus is on cardiovascular diseases

DOI: 10.4018/978-1-7998-1204-3.ch040

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(S. Muthulakshmi and K. Latha, 2012). Cardiovascular diseases take a handsome share in the chart of death threatening diseases. People are too busy in their lives; this thing added a lot in pre death (UK. Jin Yoon et al., 2010). Scenario demands a system that will have a high check over the heart's abnormality, any oddness in daily life demands a Electrocardiogram (ECG) check up (Agus Dwi Swarjaya, 2012). This demands the idea of a technology that is fast, reliable and easy to communicate. In order to satisfy this desire we have to think for a new and efficient wireless technology called Wireless Sensor Network (Ali N. Akansu et al., 2010).Moreover, this has a revolutionary impact over the various research areas, out of a large number of areas our intention is on biomedical area which cover ups a large number of diseases that comes in human lifecycle (R. Gupta and M.mitra, 2012). Electrocardiogram is one of biomedical application which is closely related to WSN (Youssouf Zatout, 2012). Rhythm of heart is a electrical activity which is governed by the traditional ECG machine by placing various electrodes (which work as sensors) that senses the electric change and the same is passed onto the machine through wires which is reportedly drawn on a graph paper in the form of waves with each represented by P,QRS, T and U (S. Karpagachelvi et al., 2010). But our scenario has changed the traditional fact by miniaturizing the size of machine and making the sensing wire free and placing it on proximity of humans by keeping in mind sometime an instant check may not report the abnormalities, so some device need to monitor the patient continuously for a period of time (Benny P.L. Lo et al.). Holter is a monitoring device which actually works in the same way as the normal ECG machines work but on a go. Usually works for 24 or 48 hours after which the stored data will be used by the doctor to report the abnormalities (Shao. Yen Tseng et al., 2009). Where if Holter instead of storing large amount of data keep on sending its acquired data continuously, but this procedure will eat up all the battery cells and Holter machines dies soon before even lasting for 24 hours because during wireless communication most of the energy is wasted. Various authors have tried to overcome this drawback to certain extent they succeeded, out of a number of reach, some have tried to work with the compression method whose job is to cut down the irrelevant data as in the previous work Discrete Cosine transformation is used which creates discrete samples, compresses them and send them wirelessly (Tellez Camiilo et al., 2009). Lack to overcome the battery life of Holter machine as Bluetooth transmission is being used and only the frequency component of ECG is taken into consideration. Our work has overpowered the drawbacks by implementing discrete wavelet transformation (DWT) as a compression method, which discretely compresses the various frequency and time samples over discrete times before compression pre-filtration is done with the help of various filters depending upon the noise ECG data contains (Li Xuemei et al., 2008). Compression only removes the unwanted or less important data from the main data but this is not enough to save energy, this puts up with the idea of optimization (Emran Mohammad Abu Anas et al.,2010).Optimization means how well the compressed data can be further enhanced without changing the meaning of the overall data (Sateh M.S. Jalaleddine et al., 1990). Optimization works on the mirror image samples, whose removal will not lead the data in unwanted state or non-meaningful (Jos Hulzink et al.,2011). Great optimization techniques have been proposed out of large category some of them are listed below-Genetic algorithm, Particle swarm optimization, Ant colony optimization etc (Rozeha A. Roshid et al., 2008).

Genetic algorithm basically gives a simple solution termed as chromosome, where its collection called as population (Hossein Mamaghanian et al., 2011). Each chromosome undergoes a process of fitness evaluation in order to check the suitability of the candidate solution for the problem. Some chromosomes undergo the process of crossover as chromosomes exchange their values or genes, where some undergo mutation where an old value is replaced by new value. Chromosomes that undergo crossover process

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and mutation process is based upon their crossover rate and mutation rates respectively (Manpreet Kaur et al., 2011). Generation puts a boundary over the evaluation of chromosome that passes through a different above mentioned processes.

Particle swarm optimization gives us a solution as particle; each particle is set to move in a space where it compares its fitness with neighbours (Mehdia Ajana EI Khaddar et al., 2012). Each particle carries two things one is the fitness value and other is its velocity. Fitness factor has two values one the value of particle it has reached so far (pbest) and second is the global value (gbest) that is decided by Particle swarm optimizer and broadcast the value to each particle in the space or population, sometime local value is broadcasted based upon topological location. Every iteration bring up with a updated value that the particle will carry, updated values include the pbest, gbest and velocities of respective particles. Iterations and velocity (Vmax) are predefined parameters.

Compression and optimization have reduced the sample rate in order to match the sample rate media for transmission is selected that is Zigbee Standard is designed and named by IEEE as 802.15.4 which actually works at low cost and low power wireless communication (Guojun Li et al., 2012). Zigbee works on the radio frequency bandwidth of 2.4 GHz / 868MHz/915 MHz. India is specified to use the 2.4 GHz range with 16 channels available with data rate of 250kb/s (P. Manimegalai et al.,2012) .Zigbee is the proprietary standard used in Wireless Communication or more precisely in Wireless Sensor communication. Zigbee operates with low power best paired with Holter machine; this in totality results in saving the battery power and allows battery to stay for few more hours.

PREVIOUS METHODS

Compression and transmission plays a pivotal and subsequent role in the field of health monitoring. Death chart grabs our attention as there is vigorous rise, due to the deadly chronic diseases that are prevailing around the humans. Among the huge number of diseases cardiac diseases are the one that need immediate attention. Cardiac diseases are visualized by the fact of Electrocardiogram (ECG) data which represents the usual heart activity of the humans in a wave cycle representing peaks as P, QRS, T and U. Various devices have been proposed but how efficient these are still a question. Efficiency depends upon the concept of data acquisition and its transmission over terrains of range while dealing with the energy optimization. Energy optimization is considered at both the hardware level and software level, there is great revolution in the field of hardware for saving energy as it gives only 30% of its share in energy saving cycle, while the rest 70% is maintained by the software part which is still in research. So far numerous software techniques have been implemented and their results have shown some satisfaction. Software part is broadly categorized as 1) Compression Techniques, 2) Optimization Techniques and 3) Transmission techniques. Over the years of research a heap of compression techniques have been proposed and modified timely. Compressing the ECG data of the patient's heart, gives us a picture to rectify the symptoms of the particular cardiac disease a patient is going through so that on time help can be availed. In our system a remarkable attention has been paid to work out with the prominent features of ECG. Discrete cosine transformation and LZW hybrid scheme has been proposed by Bo Yu et al. (2010) in their research paper. Drawbacks are as considered hybrid approach only works in the time domain of ECG and consumes more energy with lesser data rate during its transmission as media used is Bluetooth. A low power wearable cardiac device has been proposed by Maneesha V. Ramesh et al. (2012) in their research paper. Drawback different set of modes have been proposed, switching among

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

different modes is time consuming and non-efficient as energy is wasted unnecessarily while transition happens. Our proposed system is far away from the fact of mode changer and attention has been paid to the compression method's reliability and transmission media for communication. An adaptive cosine Fourier Transform approach is proposed by Patompon Chaiyakum and Paramote WardKein (2012) in their paper. Drawback is neglecting the period can put the ECG data in an unrecognizable form as time is one of the prominent factor in ECG. An optimal thresholding algorithm for denoising ECG has been proposed by Sami M. Isa et al. (2011) in their paper. Drawback although thresholding is good concept but when limited to a small section of data samples if samples are in large number it fails deliberately resulting ECG in a fix. In our proposed system has nothing to do with the thresholding concept as Samples are pre de-noised using certain set of filters which makes the compression task quite easy.

A new algorithm has been proposed by Anubhuti Khare et al. (2011) to compress the lossy ECG data in their research paper. Drawback is considering the lossy nature of ECG certain hidden corruptive signals have been neglected, and will be modified or compressed with the technique that has been used. Our proposed system deals with the hidden corruptive signals produced due to the body artifacts, and are removed quietly before compression. Wearable wireless sensor equipment has been designed to compute oxygen saturation and Heart rate and sending it on a wireless media as presented by Cristian Rotaria and Vasile Mante (2012) in their paper. Drawback of proposed system wireless media used is access point to send data over Wi-Fi, it eats up the battery soon with lesser amount of data travelled in totality. Our proposed system has tried to rectify this flaw by using another set of wireless media, which is capable of sending data efficiently with lesser energy consumption and good data rate. New compression technique has been proposed by Mohammad Abo. Zahhad et al. (2012) in which data is considered in 2D format for compression, with attention paid over the single part of ECG. Drawback of the proposed scheme is it only works for images not for digital signals in 1D, secondly it has considered only the QRS factor among the number of factors that represents ECG in more appropriate way and thirdly conversion of 1D data to 2D consumes more energy and space. Mostly the data acquisition equipments are liable to collect data in digital form; this is taken into consideration in our proposed system, working with digital signals is bit easy as compared to analog signals, whose handling is a difficult task. Optimization techniques followed by compression techniques are quite energy saving schemes.

Optimal algorithm has been designed by Grisha Garg et al. (2010) which are based upon the various thresholding concepts added with the DWT compression method shown in their research paper. Drawback in the in this chapter main focus is on the various thresholding techniques instead the kind of DWT scheme is used, thresholding adds a bit difficulty in determining the exact value for each set of thresholding schemes. Our in this chapter has nothing to do with thresholding as sample space is large but instead of thresholding concept optimization techniques have been used to remove redundancy in the samples. Ubiquitous health care monitoring equipment has been designed by Wan. Young et al. (2008) for evaluation of ECG, Accelerometer and SpO_2 presented in their research paper. Major flaw in the proposed system is that acquired data is transmitted directly without any set of compression or optimization technique, this can increment the energy and space usage during transmission. In our proposed system all these flaws have been rectified by proper usage of compression, optimization and transmission techniques. This wholesome reduces a remarkable amount of energy usage that is consumed during transmission of the data after specific steps of denoising, compression and optimization. Health decision support system is designed by Sing. Hui Toh et al. (2009) for monitoring biomedical signals as shown in their research paper. Major flaw in the proposed scheme is type of transmission media used; CDMA consumes more energy when uncompressed data travels from patient's cell phone to server where this is

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

analyzed with the preset of values. Two set of communication medias have been used that will consume more power than single media and this kind of system is feasible in local domain as its global application can lead this into complications. Wearable equipment is designed by Antonio J.Jara et al. (2012) for continuous data transmission via a Bluetooth medium as shown in their respective work. Drawback can be seen in the flow of continuous data over a media that provides lesser data rate, range and consumes more energy, for one time this can be ignored but the continuous flow of data from a wearable device can end up very early as continuous monitoring leads to the more usage of battery and the device can die soon. In our proposed system samples have been evaluated over discrete levels and treated separately over discrete times as this in whole will save the energy consumption and device can work for few more hours. Whereas the transmission media is concerned optimized wireless technology is preferred over Bluetooth for sending and receiving the discrete compressed packets.

Dealing with the redundancy found in the ECG coefficients a new algorithm has been proposed by S. Akhter and M.A. Haque (2010) entitled in their research paper. Disadvantage algorithm works only in time domain but our proposed system has taken this into consideration. ECG is represented in two broader domains as time domain and amplitude domain. ECG data holds the various corruptive signals which hide its functionality, steps to be considered in order to remove those corruptive signals produced by the various movements of the body before the compression. Optimal way of recognizing the various abnormalities in ECG data using Neuro-Fuzzy approach shown by P.D. Khandait et al. (2012) in their research paper. In this paper major drawback is more energy, time and space consuming with less concentration on the artifacts that causes abnormalities every set of abnormality whose negligence cannot be afforded is not recognized by Neuro-Fuzzy technique. In our in this chapter we have statically mentioned the major abnormalities and removed them under a course of steps this saves time, energy and space, precisely design comes up with the data that is important and make it ready for transmission after removing set of abnormality as the received data is an impression for the doctor to find out the oddness so as to reach the patient instantly and timely.

CHAPTER'S HIGHLIGHTS

Electrocardiogram (ECG) data can be compressed by number of techniques but technique mentioned in the base paper is DCT (Discrete Cosine Transform) followed by LZW (Emran Mohammad Abu Anas et al.,2010) and compressed data is sent over the Bluetooth environment, as DCT is restricted to the certain compressed ratio and it blocks some of the important features of the signal and another drawback of the DCT (Anubhuti Khare et al., 2011)is that it represents the signal in the frequency domain but ECG composed of two domains as frequency and time domains. Both domains are equally important to extract the right meaning of the ECG data, for this Discrete wavelet transform (DWT) is used to represents the signal in the frequency and time domain. After compression the second important thing is the transmission of data, whereas in the base paper transmission is done over the Bluetooth media. Medium that has been chosen seems unreliable and more power consumer which results in early dying of the Holter's battery, these drawbacks have been corrected in our chapter with full attention has been paid to the energy utilization. Wavelet transformation compression technique and Zigbee environment overcomes the drawbacks of DCT (Patompon Chaiyakum & Paramote WardKein, 2012) and Bluetooth respectively. Some set of add-on is done in the proposed scheme that is optimization technique, which further reduces the number of samples to be sent over the wireless media.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

NEED AND SIGNIFICANCE OF THIS CHAPTER

Need of the in this chapter is illustrated by the Related work as said that ECG data that is evaluated continuously while patient is involved in his daily activities, and secondly its transmission consumes more energy than energy consumed during its compression. Continuous ECG (Uk.Jin Yoon et al., 2010) will be compressed in discrete form under wavelet transformation and transmitted over the Zigbee IEEE802.15.4 protocol and reaches the Doctor or care taker for immediate assistance. Continuous monitoring of data and its transmission without compression can lead the device out of power after few hours but if we correct the transmission by just sending the compressed data over time-domain as a constraint and better protocol for transmission, by this power can be saved and the device can operate for few more hours under set of various environmental conditions. Whereas the significance is concerned it is illustrated by the fact that energy is being saved on account of Compression method and transmission is made better by use of Zigbee IEEE802.15.4 (Wan.Young Chung et al., 2008) protocol as it consumes less power for transmission, easy to operate than Bluetooth (Antonio J. Jara et al., 2012) along with these techniques some optimization techniques have been proposed which further reduces the number of samples to be sent. Results signify the better working of the proposed system under set of parameters as energy. IEEE802.15.4 has a peculiar significance that although its data rate is less than the Bluetooth but consumes less power while transmitting the data as compared to the Bluetooth wireless technology.

EXPERIMENTAL DESIGN

In this paper we have proposed a approach to compress the ECG data after removal of noise from the raw ECG data collected which is obtained from the real time simulator through Virtual Labs (Vlabs) and then passing it through optimizers and then the optimized data is tested over the Zigbee (802.15.4) Environment. Although some performance metrics have been used for evaluation. Once the different ECG data (normal, muscle, respiration and baseline wandering) are obtained from the simulator these undergo a noise removal process respectively at very first step of the proposed signal processing algorithm. Three basic types of noises such as muscle artifact, respiration artifact and baseline wandering will be eliminated which can be probably caused by muscle movement, breathing problem and in contact with high voltage respectively. The first step is to eliminate the baseline drift/wandering (Manpreet Kaur et al., 2011) which is a type of noise got generated during the signal recording process. Baseline drift will be eliminated using a smooth filter.

The next step is to remove the muscle artifact using median Filter. After removing the muscle artifact, the next objective is to remove the spikes or respiration artifact from the ECG data that arises because of any voltage fluctuation using average Filter. Various metrics have been evaluated before and after filtration process using equations 1, 2 and 3.

Then the single-level one-dimensional wavelet decomposition is performed on the ECG signal, which decomposes the signal using discrete wavelet, DWT compression usually compresses the whole received pre-processed signal and transforms it in two components one is approximation coefficient and other is detail coefficient. Coefficients designed above envelopes vital information that is passed on for further processing respectively. Along with Compression some metrics have been evaluated like PSNR (GirishaGarg et al., 2010), MSE (Sani M. Isa et al., 2011) and PRD (Mohammed Abo.Zahhad et al., 2012) by following equations 1,2 and 3.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$PSNR = 10 * \log_{10}\left(\frac{255 * 255}{MSE}\right) \tag{1}$$

$$MSE = \frac{1}{n} * \left(\sum_{i=0}^{n} (x[i] - x'[i])^2\right) \tag{2}$$

$$PRD = \sqrt{\frac{\sum_{n-0}^{n-1} (x[i] - x'[i])^2}{\sum_{n-0}^{n-1} (x[i])^2}} * 100 \tag{3}$$

x[i]- Original ECG data, x'[i]- Transformed ECG data, n- Number of samples.

Above metrics are evaluated at two stages that is at the former stage of compression and secondly at the latter stage of compression. The decomposed signal is passed to the particle swarm optimization and genetic algorithm for the signal optimization process. The ECG signal is optimized using the above two algorithm to get the best ECG signal out of the obtained ECG signal.

The detail of our execution of GA has been described as follow;

Genetic Algorithm Procedure

Input: Compressed ECG data
Output: Optimized ECG data
Step 0: attribute the various variables (population size, crossover rate, mutation rate and utmost number of generations)
Step 1: initialize population randomly as (G(0))
Step 2: calculate present population (find the fitness of chromosomes/genes)
Step 3: while (condition not met) do [step 4-8]
Step 4: select G(u) from G(u+1) [execute selection process]
Step 5 again combine G(t)[do mutation and crossover]
Step 6: caluculate present population (find the fitness of chromosomes/genes).
Step 7: u=u+1
Step 8: go to step 3 repeat the process again

The detail of our execution of PSO has been described as follow;

Particle Swarm Optimization Procedure

Input: Compressed ECG data
Output: Optimized ECG data
Step 1: For each bird/value
Step 2: attribute the values/birds
Step 3: End

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Step 4: Do
Step 5: For each bird/value
Step 6: apply fitness check
Step 7: If resulted fitness is enhanced than pBest in history
Step 8: set pBest equal to current value
Step 9: End
Step 10: Choose the pbest among all the birds/values assign it as the gBest for all the birds repeatedly.
Step 11: evaluate bird velocity with equation (a)
Step 12: Update bird position with equation (b)
Step 13: End

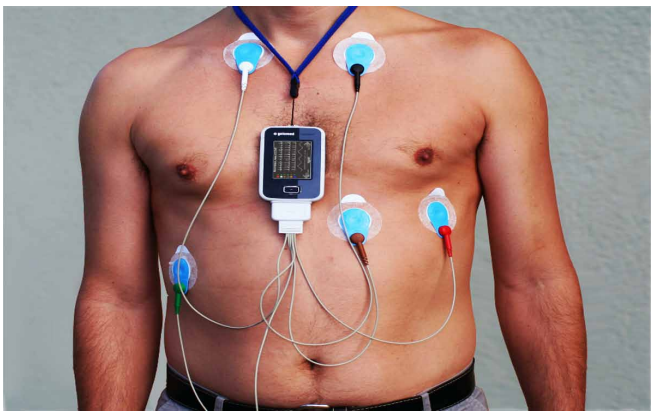
$$\mathrm{a[i]} = \mathrm{a[i]} + \mathrm{d\,1} * \mathrm{rand}() * (\mathrm{pb[i]} - \mathrm{pre[i]}) + \mathrm{d\,2} * \mathrm{rand}() * (\mathrm{gb[i]} - \mathrm{pb[i]}) \tag{4}$$

$$\mathrm{pre[i]} = p\,\mathrm{re[i]} + a\mathrm{[i]} \tag{5}$$

a[i] - particle velocity; pre[i] -current solution that particle holds randomly; pb[i] and gb[i]-local and global best fitness values respectively for each particle; rand ()- random number (0,1), d1, d2 -are learning factors [0 2].

Energy is calculated for both of the optimization processes that is for Genetic algorithm and Particle Swarm Optimization. The one with the lesser energy will be allowed to forward its Optimized compressed ECG data (noise free) for transmission to the medical database server for further analysis by doctor via Zigbee environment experimented on a hardware module. Zigbee consumes lesser power and work without any fatal error. This allows ZigBee networks to form ad-hoc, with no control from the centralized department or high-power transmitter/receiver able to reach all of the devices. Received ECG data is the optimized data along with position where the best sample is residing is used for decompression process. The examined result and all the vital information is later saved to the medical database along with the received ECG signal under the patient ID from where it was received earlier, by this immediate medical assistance is stipulated to the targeted patient and further patient is under regular surveillance (Figure 1).

Figure 1. Flowchart of experimental design



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Step 1: Different types of ECG data are collected from the online virtual lab (Vlab) site simulated in real Time

Step 2: Collected ECG data are termed as normal, muscle artifact, baseline wandering and respiration artifact.

Step 3: Extracted data is passed on to different filters for removal of noise based upon the type of collected data filter is used.

Step 4: Muscle artifact data is passed to the median filter for removal of spikes, baseline wandering data is passed to smoothing filter and respiration artifact data is passed to average filter respectively.

Step 5: Performance metrics have been evaluated before and after filtration.

Step 6: Pre-processed data is ready to undergo a wavelet Transformation DWT

Step 7: Performance metrics have been evaluated before and after compression according to equation number 1,2 and 3 respectively.

Step 8: while (iteration terminated)

Do (Step 8)

Step 9: Compressed ECG data passed through optimizers named as Genetic algorithm and Particle swarm optimization.

Step 10: Best optimized ECG data obtained based upon energy consumed by both optimizers, one with less is preferred

Step 11: Optimized compressed ECG data ready for transmission over Zigbee environment.

Step 12: Received ECG data analyzed and stored in medical database for further analysis

Step 13: End.

SIMULATION RESULTS

In this research, a novel heart beat signal propagation and computation model is developed for the patients using body wearable sensors for heart rate monitoring (also known as Holter) connected through Zigbee Environment with the medical database. These kinds of scenarios are usually designed for the special medical environment where patients are monitored on the regular basis to avoid the critical health hazard situations. In this research project, the aim is to improve the ECG signal propagation process by making the whole process energy efficient. The energy consumption can be minimized by optimizing the signal, which will yield out a small-sized equally detailed optimized ECG signals. This optimized ECG signal can be propagated on the ZigBee environment by consuming less energy than the non-optimized signal, for sake of easiness hardware Zigbee module is designed for communication. Also, one more aim of this research is to compute the energy consumption by Genetic Algorithm and Particle Swarm Optimization on the Holter (Body Wearable Heart Rate Recording Wireless Sensor). The signal with the lower energy consumption will be propagated on the ZigBee environment and other complex signal consuming more energy will be discarded.

Various patients have been put on test and evaluation under same set of conditions and various set of performance metrics have been evaluated as- peak signal to noise ratio, mean square error and percentage root mean square difference, all these when taken into consideration have come up with new results that shows remarkable difference among the values before and after filtration and compression respectively based upon different filter size. Tabular representation shown in Table 1 shows that peak signal to noise ratio is more after compression than before compression, whereas mean square error is

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

less after compression than before and same is with the percentage root mean square value which shows a decrement in its value after compression than before compression. Simulated results show that error or corruptive signals have been removed during compression before the signal is ready for transmission.

Table 1 shows the various values of performance metrics before filtration of noisy ECG data with different window sizes (5, 6, 7, 8, 9) respectively. Comparison represents that in case of normal data the value of PSNR is infinite for respective window size, value for MSE is zero and value for PRD is 100 for respective window size, this shows that the normal data is accurate and further comparisons can be drawn from this normal ECG data. Various values of Peak signal to noise ratio, mean square error and percentage root mean square difference for the muscle artifact, respiration artifact and baseline wandering ECG data are listed below with respect to the normal ECG data.

Record shows that values differ from their parent values; apart from their parent values they have values that are different, which gives us a sign that may be some noise is present which needs certain methods for their removal.

The results in the below Table 2 has proved that after filtration of noise from muscle artifact, respiration artifact and baseline wandering ECG data various performance metrics have changed where PSNR has decreased constantly and the same fall is seen in PRD Values and but there is a increase in MSE values with different window sizes respectively which illustrates that there is an noise/error in the signal has with respect to the parent data. Calculated values in Table 2 also represent the values before compression of ECG data (Muscle artifact, respiration artifact and baseline wandering free).

Table 3 shows the values of various performance metrics after the compression process which shows that there is an increase in the PSNR, same increase in seen in PRD values but there is a fall in MSE values with different window sizes respectively which shows that error has been removed now the data is error free as shown in tabular form.

In Table 4 particle swarm optimization outperformed Genetic Algorithm. PSO has taken almost less than half energy than GA for the ECG signal optimization. Each patient's ECG data is examined in this project is of the length 40 seconds and has recorded on the 100 samples per second as a sampling rate.

This research project has been developed along with a GUI using the graphical user interface developer called GUIDE in the MATLAB software. The GUI contains various push buttons, text fields and signal plot axis. The buttons are used to process all of the tasks listed in the project one after one in a particular sequence. The processes listed to run on the buttons are main simulation, import signal (normal, muscle artifact, respiration artifact and baseline wandering ECG data), filtration (median, average and smooth filter), DWT filter, optimization (GA & PSO) and Comparison (GA and PSO). The general GUI is shown in the Figure 2.

Table 1. Recorded performance parameters (statistical errors) before filtration

Window Size	Normal ECG Data			Muscle Artifact			Respiration Artifact			Baseline Wandering		
	PSNR	MSE	PRD	PSNR	MSE	PRD	PSNR	MSE	PRD	PSNR	MSE	PRD
5	Inf	0	100	60.39	0.05	93.13	61.5	0.04	96.6	61.4	0.04	98.3
6	Inf	0	100	60.39	0.05	93.13	61.5	0.04	96.6	61.4	0.04	98.3
7	Inf	0	100	60.39	0.05	93.13	61.5	0.04	96.6	61.4	0.04	98.3
8	Inf	0	100	60.39	0.05	93.13	61.5	0.04	96.6	61.4	0.04	98.3
9	Inf	0	100	60.39	0.05	93.13	61.5	0.04	96.6	61.4	0.04	98.3

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Recorded performance parameters (statistical errors) after filtration or before compression

Window Size	Muscle Artifact			Respiration Artifact			Baseline Wandering		
	PSNR	MSE	PRD	PSNR	MSE	PRD	PSNR	MSE	PRD
5	69.7	0.006	69.4	66.0	0.016	72.55	83.47	0.0002	95.508
6	67.7	0.011	60.1	65.58	0.017	67.65	83.47	0.0002	95.508
7	66.5	0.014	55.5	65.26	0.019	63.47	74.55	0.0022	87.75
8	65.9	0.013	52.0	65.00	0.02	59.77	74.55	0.0022	87.75
9	65.4	0.018	51.0	64.79	0.021	56.40	70.58	0.0056	80.72

Table 3. Recorded performance parameters (statistical errors) after compression

Window Size	Muscle Artifact			Respiration Artifact			Baseline Wandering		
	PSNR	MSE	PRD	PSNR	MSE	PRD	PSNR	MSE	PRD
5	63.4	0.02	91.92	63.83	0.021	93.87	61.66	0.044	85.38
6	64.69	0.022	93.98	64.31	0.024	94.99	61.66	0.044	85.38
7	65.73	0.01	95.52	64.72	0.02	95.54	62.72	0.03	89.08
8	66.36	0.01	96.44	65.10	0.020	95.95	62.72	0.03	89.08
9	66.74	0.013	96.60	65.48	0.018	96.30	63.77	0.02	92.26

Table 4. Energy consumption by PSO, Genetic Algorithm and ZigBee -Environment

Sample Types	Energy Computed by GA	Energy Computed by PSO
Normal ECG data	49.03	27.42
Muscle artifact ECG data	50.49	26.98
Respiration artifact ECG data	48.10	37.75
Baseline wandering ECG data	49.04	25.22

Figure 2 and 3 shows the imported ECG data from the Virtual Labs. The various noises in the signal are removed is the step after plotting the original signal as shown in Figure 3 and 4.

After the respiration artifact and muscle artifact are being removed a significant change occurs in the signal structure shown in the Figure 5.

Noise produced by the baseline wandering is being removed such that the structure of the signal is changed as shown in Figure 6.

After the process of filtration process each ECG data is passed through compression module which in turn decreases the number of samples to be sent as shown in Figure 7 and 8 respectively.

After the compression process, compressed ECG data still contains certain mirror samples which further are removed with the optimization process under genetic algorithm and particle swarm optimization, individual algorithm is checked for the energy consumed by them during optimization process and one with lesser energy is preferred as shown in Figure 9 and 10 respectively then this data is allowed to be sent over wireless environment.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. GUI interface figure

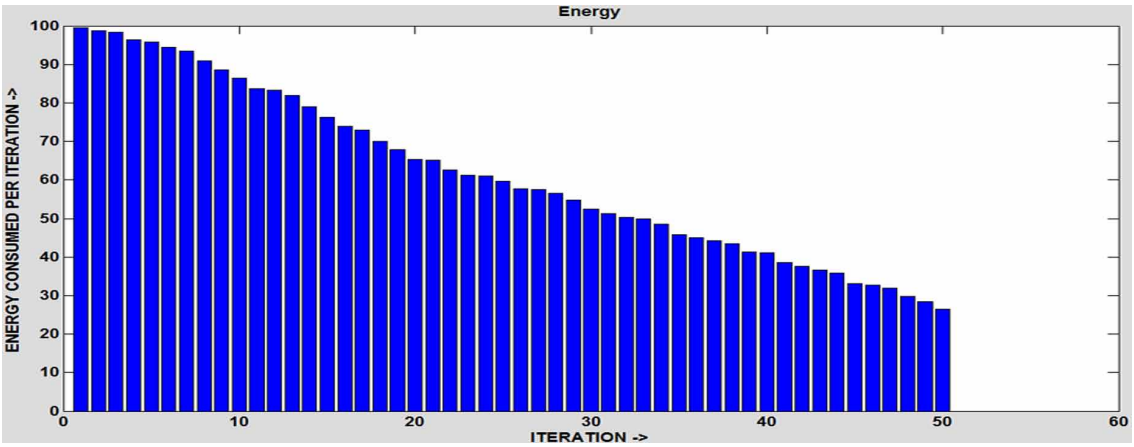


Figure 3. Displaying normal and muscle artifact ECG signal

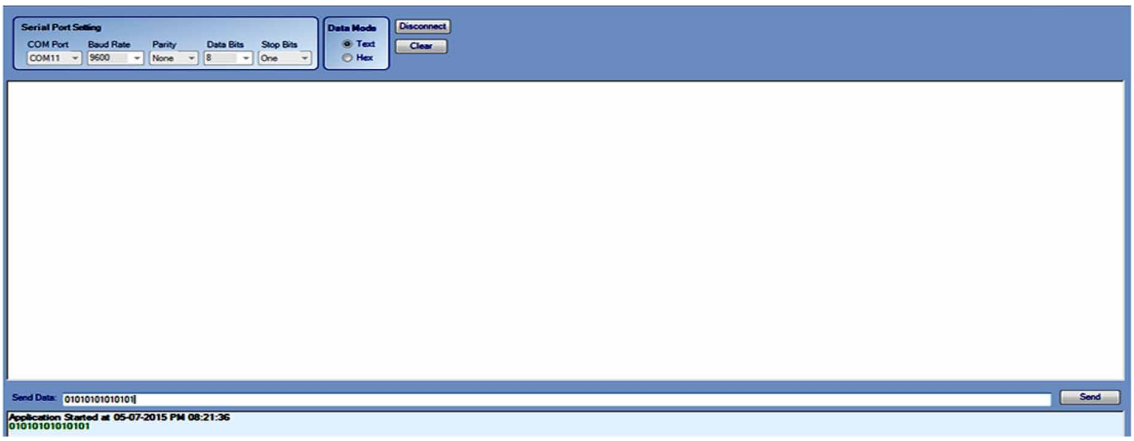
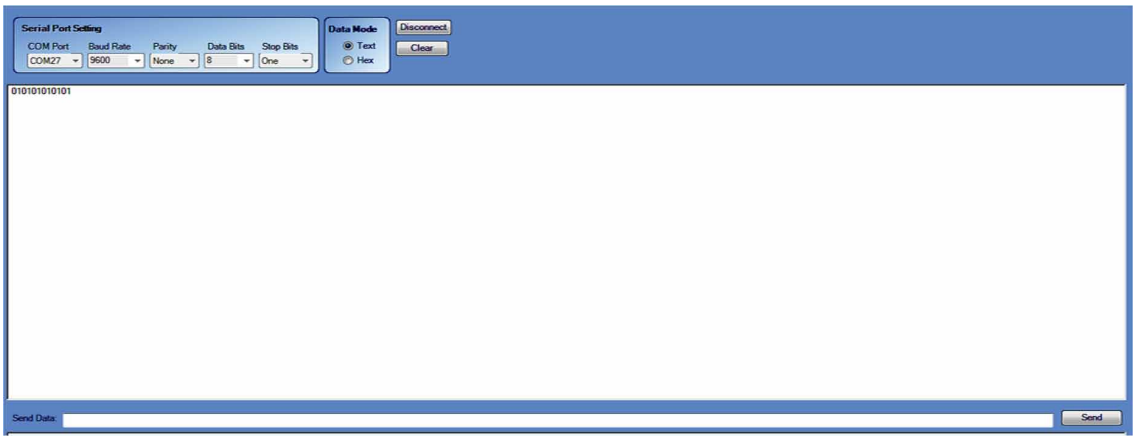


Figure 4. Displaying respiration artifact and baseline wandering in ECG signal



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Plot of respiration, muscle artifact ECG data after application of average, median filter respectively

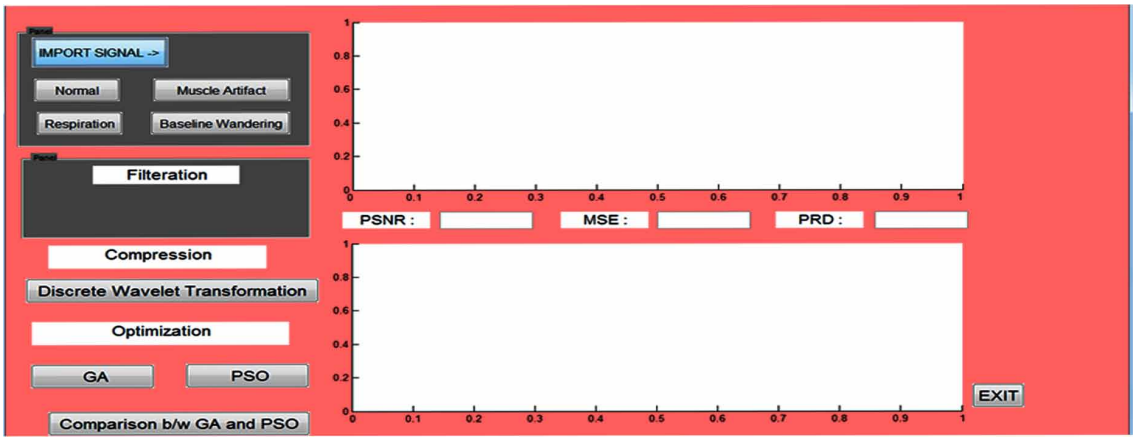


Figure 6. Plot of baseline wandering ECG data after application of smoothing filter

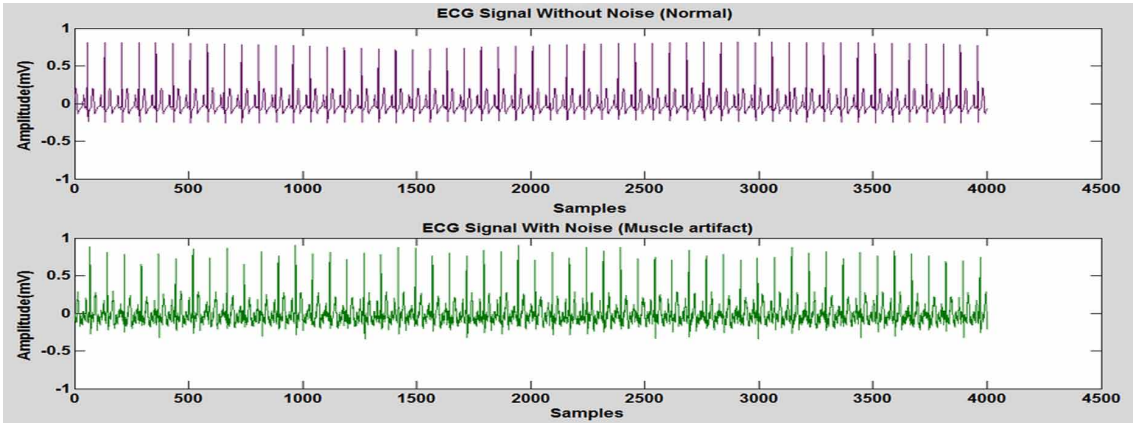
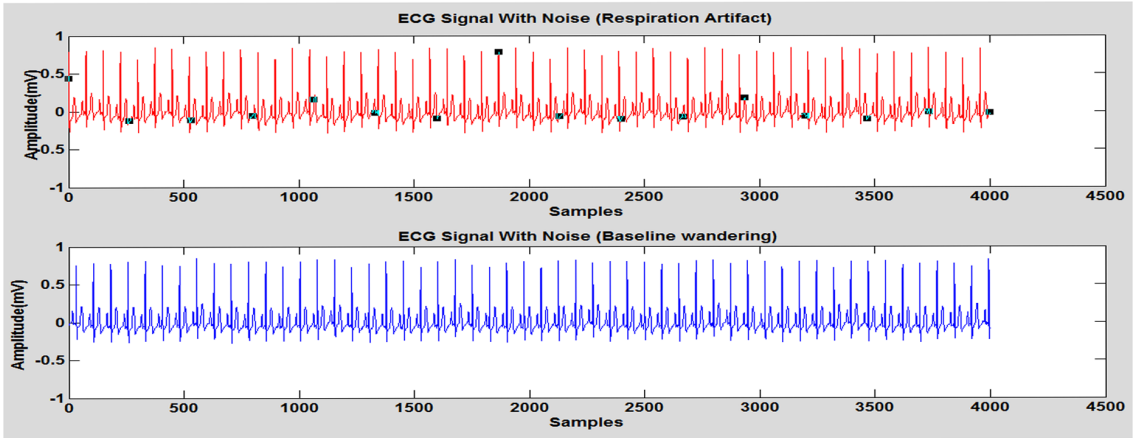


Figure 7. Compression of normal and muscle artifact free ECG data respectively



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Compression of respiration artifact and baseline wandering free ECG data respectively

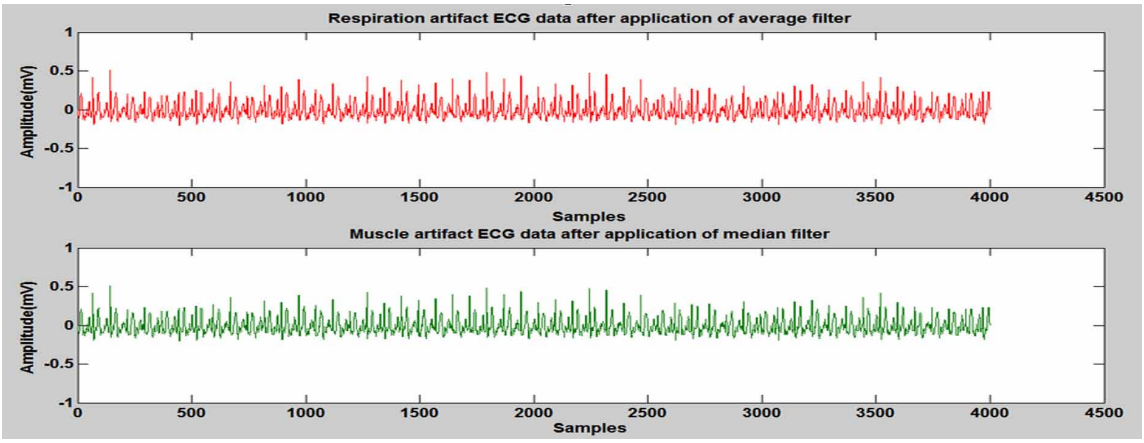


Figure 9. Energy consumption of muscle artifact ECG data under GA

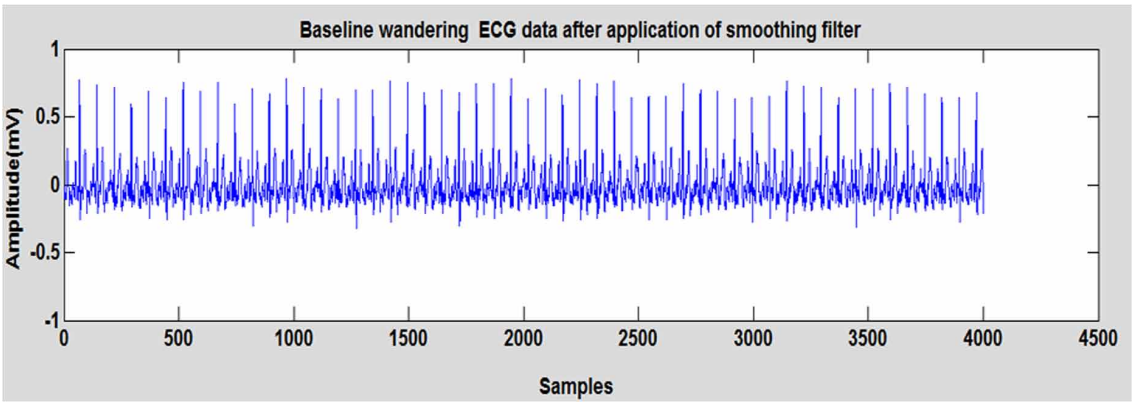
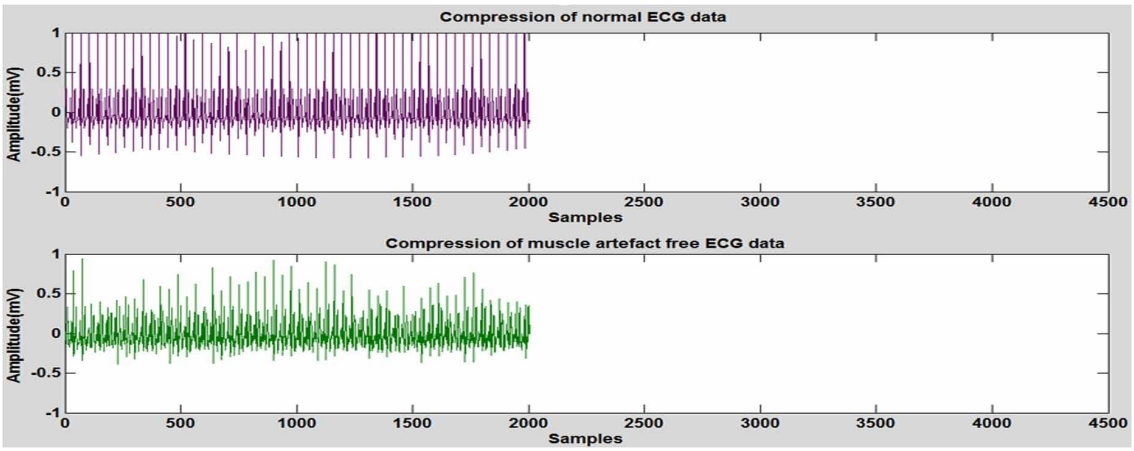


Figure 10. Energy consumption of muscle artifact ECG data under PSO



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Optimized data received from the algorithm that consumes lesser energy, in our scenario which shows that PSO is dominating GA. Former data is ready for transmission over the Zigbee environment, in order to ease the transmission a hardware Zigbee module is designed which a circuit board embedded with transformer, followed by microcontroller, diodes, resisters, capacitors, Zigbee transmission mote, MAX-232 cable to connect it with the Personal computer, and the whole system takes the battery and system of personal computer for its proper working. Figures 11 and 12 shows the transmission process over Zigbee hardware module transmitted to and from MATLAB with the set of various instructions.

Figure 11. Transmission of optimized data over Zigbee module

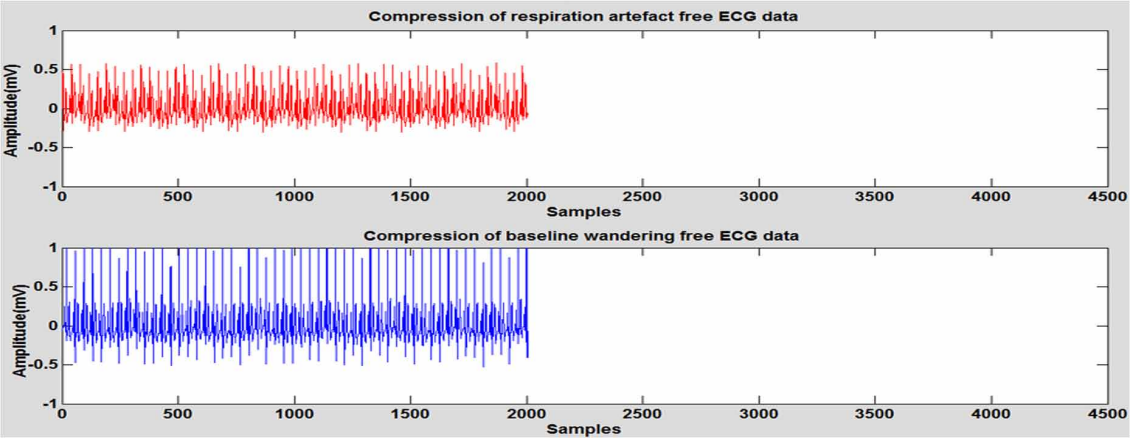
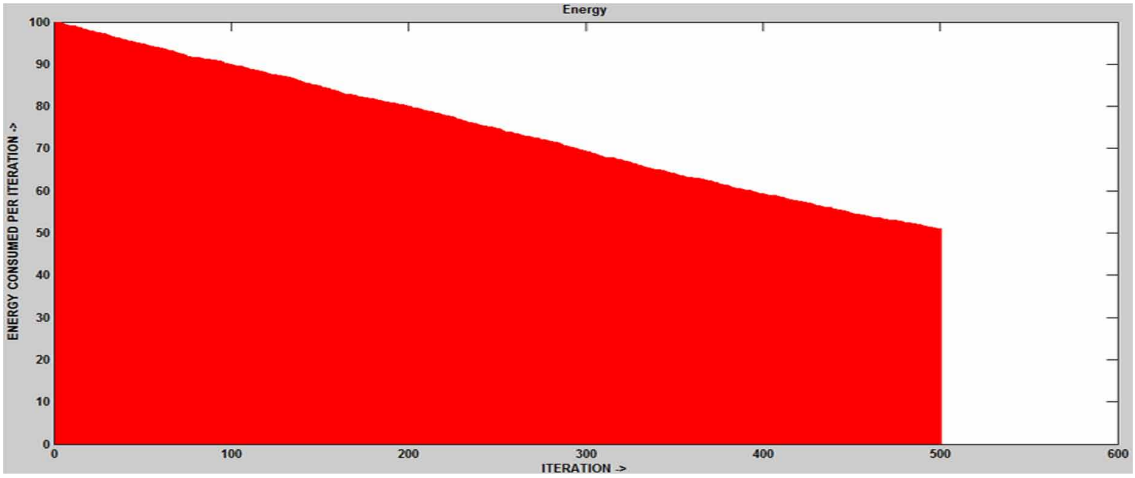


Figure 12. Receiving of optimized data over Zigbee module



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

There is a future scope that further takes this to another level as the combination of various compression algorithms can be implemented on the various types of ECG data. Pre-filtering of the noisy ECG data can be done by the applications of various hybrid filters further work can be considered based upon the transmission process of wireless communication. Moreover, bandwidth dedication, channel allocation and delay related factors are some of the issues on which the work can be performed, as this will not only enhance the data transmission but will allow the device to send the more accurate data on-demand. In addition to this, certain other artifacts that can be identified with the help of the ECG transmitted will be an icing on the cake. What's more, work can be further carried forward by considering the various uproars near the patient who is under surveillance. Further, this chapter gives someone a platform to deal with the high prioritised discipline that is health, which nowadays demands a lot more attention than any other field of the education, as this field highly represents the way to deal with all the detrimental problems and how to curb with those diseases by initiating certain steps before it is too late for the patient, who actually needs this kind of help for themselves.

CONCLUSION

In this research project, we have proposed an effective ECG signal optimization and propagation framework using particle swarm optimization (PSO), Genetic Algorithm (GA), ZigBee wireless environment, discrete wavelet transform, noise removal filters and signal drift removal filter techniques and technologies. All of the technologies have been combined on one platform to increase the performance of the ECG signal recording using body wearable wireless sensors for ECG monitoring and low energy propagation of the ECG signal by compressing the signal using optimization techniques irrespective of the loss of ECG signal details. When compared with the base paper which includes only DWT compression over different mode consumes more energy when devices switch from one mode to other. In the base paper [1] compressed data was allowed to traverse through Bluetooth technology which is not efficient and consumes lot of energy and data packets are of low rate which is not feasible. Above two problems have been rectified in the proposed chapter with due concentration over the Energy consumption as there is no mode change it operates continuously but sends data discretely based upon the DWT compression technique. Maximum energy is consumed during transmission so Zigbee environment is preferred over Bluetooth to send the data with data rate higher than Bluetooth. Proposed paper do include another set of energy saving technique that is optimization techniques which are missing in the base paper, this overall saves more energy while data that is meant for sending is lesser in amount as unwanted or repeated data will be flushed away by PSO and GA . Out of two one that has consumed significantly lower energy than the other for the process of ECG signal optimization will be traversed over communication channel. Whereas it is also recorded that ZigBee environment has also consumed significantly lower amount of energy than the non optimized ECG signal propagation environment. Transferred data is being stored in a database for further analysis.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Akansu, A. N., Serdijn, W. A., & Selesnick, I. W. (2010). *Emerging applications of wavelets: A review*. Elsevier.

Akhter, S., & Haque, M. A. (2010, August). Ecg Comptression Using Run Length Encoding. In *18th European Signal Processing Conference EUSIPCO-2010* (pp. 1645-1649).

Camiilo, T., Oscar, R., & Carlos, L. (2009). *Biomedical Signal Monitoring Using Wireless Sensor Networks*. IEEE.

Chaiyakum & Wardkein. (2012, May). A New ECG Compression by Using Adaptive Cosine Fourier Transform. In Electrical Engineering/Electronics, Computer, Telecommunication and Information Technology (ECTI.CON) (pp. 1-4).

Chung, L., & Jung. (2008, August). A Wireless Sensor Network Compatible Wearable U.Healthcare Monitoring System Using Integrated ECG, Accelerometer and SPO2. In *30th Annual International IEEE EMBS conference* (pp. 1529-1532).

Garg, S. Gupta, & Mittal. (2010). Optimal Algorithm for ECG Denoising using Discrete Wavelet Transforms. IEEE.

Gupta, R., & Mitra, M. (2012, December). An ECG compression technique for telecardiology application. In *India Conference* (p. 4). IEEE.

Hossain, Afran, & Sayed. (2010, December). Compression of ECG signals exploiting correlation between ECG cycles. In *6th International Conference on Electrical and Computer Engineering ICECE 2010, Dhaka, Bangladesh* (pp. 622-625).

Hulzink, Konijnenburg, Ashouei, Breeschoten, Berset, Huisken, … Van Ginderdeuren. (2011, December). An Ultra Low Energy Biomedical Signal Processing System Operatng at Near Threshold. *IEEE Transactions on Biomedical Circuits and Systems, 5*(6), 546-554.

Isa, Noviyanto, & Arymurthy. (2011). Optimal Selection of Wavelet Thresholding Algorithm for ECG signal denoinsing. In ICACSIS (pp. 365-370).

Jalaleddine, Hutchens, Strattan, & Coberly. (1990, April). ECG Data Compression Techniques. A Unified Approach. *IEEE Transactions on Bio-Medical Engineering*, *37*(4), 329–343.

Jara. (2012, July). Evaluation of Bluetooth Low Energy capabilities for continuous data transmission from a wearable electrocardiogram. In Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS) (pp. 912-917).

Karpagachelvi, Arthanari, & Sivakumar. (2010, April). ECG Feature Extraction Techniques-A Survey Approach. *International Journal of Computer Science and Information Security*, *8*(1), 76–80.

Kaur, M., Singh, B., Ubhi, & Rani. (2011). Digital Filtration of ECG Signals for Removal of Baseline Drift. In *International Conference on Telecommunication Technology and Applications Proc. of CSIT.* IACSIT Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Khaddar. (2012, May). Emerging Wireless Technologies In E-Health. In Multimedia Computing and System (ICMCS), (pp. 440-445).

Khandait, Bawane, Limaye, & Khandait. (2012, April). Efficient ECG Abnormalities Recognition Using Neuro-Fuzzy Approach. International Journal of Advanced Research in Computer Science and Software Engineering, 2, 402-407.

Khare, Saxena, & Nerkar. (2011). ECG Data Compression Using DWT. *International Journal of Engineering and Advanced Technology*, *1*(1), 11–13. doi:10.1109/HealthCom.2012.6379443

Li, J., & Li. (2008, May). Home healthcare platform based on Wireless Sensor Networks. In *5th International Conference on Informational Technology and Application in Biomedicine*, China (pp. 263-266). doi:10.1109/LATINCOM.2009.5305161

Li. (2012, May). Genetic Particle Filtering for Denoising of ECG Corrupted by Muscle Artifacts. In *Natural Computation (ICNC)* (pp. 562-565).

Lo, Thiemjarus, King, & Yang. (n.d.). *Body Sensor Network – A Wireless Sensor Platform For Pervasive Healthcare Monitoring.* Academic Press.

Mamaghanian, Khaled, Atienza, & Vandergheynst. (2011, September). Compressed Sensing for Real Time Energy Efficient ECG Compression on Wireless Body Sensor Nodes. *IEEE Transactions on Biomedical Engineering, 58*(9), 2457-2466.

Maneesha, V., Ramesh, R. G. R., & Abhishek, T. K. (2012). Low Power Intelligent Wearable Cardiac Sensor Using Discrete Wavelet Compression. In *International Conference on Advances in Mobile Network, Communication and its Application* (pp. 107-110). IEEE.

Manimegalai, P. (2012, March). Estimation of Cardiovascular Parameters from ECG and PPG Signals. In *Bioengineering Conference (NEBEC)* (pp. 237-238). 10.1109/NEBC.2012.6207051

Muthulakshmi, S., & Latha, K. (2012). Classification of ECG Waveform Using Feature Selection Algorithm. In *International Conference on Advanced Communication Control and Computing Technologies* (pp. 162-165). IEEE.

Rashid, A., & Rahim, S. & Mahalin. (2008, December). Home healthcare via Wireless Biomedical Sensor Network. In *IEEE International Rf And Microwave Conference Proceedings* (pp. 511-514). IEEE.

Rotariu & Manta. (2012). Wireless System for Remote Monitoring of Oxygen Saturation and Heart Rate. In *Proceedings of Federated Conference on Computer Science and Information System* (pp. 193-196). IEEE.

Suarjaya. (2012). A New Algorithm for Data Compression Optimization. *International Journal of Advanced Computer Science and Applications, 3*(8), 14-17. doi:10.1109/ICACCCT.2012.6320762

Toh, Do, Chung, & Lee. (2009). Health Decision Support for Biomedical Signals Monitoring System over a WSN. In *2nd International Symposium on Electronic Commerce and Security* (pp. 605-608). IEEE.

Tseng, Tsai, Lai, & Fang. (2009). *A Wireless Biomedical Sensor Network using IEEE802.15.4*. IEEE.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yoon, Hwang, Noh, Cheol, & Royoon. (2010, July). Comparison Of Cwt With Dwt For Detecting QRS Complex On Wearable Ecg Recorder. In *IEEE, International Conference on Wavelet Analysis and Pattern Recognition, Qingdao* (pp. 300-303). IEEE.

Yu, Yang, & Chong. (2010). ECG Monitoring over Bluetooth: Data Compression and Transmission. In *Communication Society WCNC Proceedings*. IEEE.

Zahhad, Ahmed, & Zakaria. (2012). Electrocardiogram Signals Compression Technique Based on QRS. Complex Estimation and 2.D DWT. In ICCIT (pp. 554-559).

Zatout, Y. (2012). Using Wireless Technologies for Healthcare Monitoring at Home: a Survey. In 14th International Conference on e.Health Networking, Applications and Services (pp. 383-386). IEEE.

This research was previously published in Computational Tools and Techniques for Biomedical Signal Processing edited by Butta Singh; pages 314-332; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 41
Electromyogram and Inertial Sensor Signal Processing in Locomotion and Transition Classification

Deepak Joshi
Indian Institute of Technology (IIT) Delhi, India

Michael E. Hahn
University of Oregon, USA

ABSTRACT

Signal processing in biomedical engineering is essentially required for classification while serving mainly two aims. The first is noise removal and the second is signal representation. Signal representation deals with transforming the signal in such a way that the signal is most informative in that particular domain for the application at hand. This chapter will describe signal processing methods like spectrogram with specific applications to locomotion and transition classification using Electromyography (EMG) data. A wavelet analysis application on foot acceleration signals for automatic identification of toe off in locomotion and the ramp transition is also shown. Finally, the performance of EMG and accelerometer performance across different time windows of a gait cycle in locomotion and transition classification is presented with an emphasis on fusing the data from both sensors for better classification.

INTRODUCTION

A signal is a single-valued representation, real or complex, of information as a function of an independent variable recorded from some sensing device. Physically, it is a form of energy being produced by a process. Various types of signals exist and the method of signal processing needed to process a signal heavily depends on the type of the signal. Broadly, there are four types of signals:

DOI: 10.4018/978-1-7998-1204-3.ch041

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. Deterministic,
2. Stochastic,
3. Chaotic, and
4. Fractal.

Deterministic signals can be represented mathematically using certain parameters and hence prediction of deterministic signal in future, based on past values, is possible. Stochastic signals are random in nature with no closed loop mathematical expression and hence are not predictable in future. Fractal signals have interesting property of scale-invariance i.e. at any level of magnification the signal looks similar in structure. Chaotic signals lie somewhere between deterministic and stochastic signals. They deemed to be deterministic in a very sensitive sense and hence are not predictable in future with reasonable accuracy. Biomedical signals are generally stochastic in nature and need signal processing methods for various applications. Biomedical signal processing in various applications serves mainly two aims. The first is noise removal and the second is signal representation. The main source of noise in biomedical signals is noise in electronic instrumentation. Other sources consist of electromagnetic interference and skin motion or physiological artifacts. For example, while recording electromyogram (EMG) nearby the chest one can observe the artifacts of electrocardiogram (ECG) in the EMG profile. There have been many provisions in hardware and experimental protocols to remove such noise and artifacts, however complete removal is not possible without using signal processing tools. Signal representation deals with transforming the signal in such a way that the signal is most informative in that particular domain for the application at hand. One such application is classification of locomotion and transitions in assistive devices where noise removal and signal representation are critical for successful, rapid and safe classification.

Locomotion classification is crucial in lower limb prosthesis control. Powered lower limb prosthesis provides sufficient torque and stability to the amputee so that they can negotiate in different terrains with reduced energy consumption. The key component for powered prosthesis to be facilitated with such features is accurate locomotion classification. Based on the classified locomotion, the artificial limb changes the damping profile to provide with net required torque output. For seamless transition in a powered prosthesis it is critical to classify the upcoming locomotion ahead of time (ideally before toe off of the transitional gait cycle) to ensure safe transition. Hence the selection of required data is critical and depends on the application at hand. For example, if someone is interested in analyzing the offline data of a cerebral palsy patient then the full gait cycle (heel strike to heel strike) may be available. However, for the application of real time actuation in prosthesis and orthosis control, the data available is limited. Hence, within this framework of classification the selection of sensors is critical. EMG represents the muscular activity and has been used widely for various applications like prosthetic control, gait analysis, disease diagnosis, and neuro-muscular coordination (Gea Drost et.al. 2006; Anjana Goen and D. C. Tiwari 2013). With the advancement in technology, miniaturized inertial sensors like accelerometers have shown promise in locomotion classification and other clinical gait analysis applications. Ambulatory monitoring of daily activities is the other area where accelerometer is of great importance as a sensing device.

EMG is the summation of electrical activity, action potential, of all the motor units recorded by the electrode. EMG signals are recorded by placing electrodes in, on, or near a muscle and amplifying the electric potential between two electrodes, that results from the generation and propagation of action potentials along muscle fibers. These muscle fibers are innervated by motor neurons from the spinal cord. The electrodes used are generally made of silver/silver chloride (Ag/AgCl) with some application of conductive gel at the skin-electrode interface. The electrodes serve as a transducer which converts

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the ionic activity of cells inside the body into the electric current outside the body. To reduce the effect of common mode electromagnetic noise the electrodes are used in differential mode. Such a differential mode also offers to measure localized activity within the muscle. After removing the noise from the EMG signal it is required to represent the EMG signal for the application in hand. Signal representation of EMG signal for various applications essentially consists of representing the signal in 1) time domain, 2) frequency domain, and 3) time-frequency domain. Time domain processing may include autocorrelation, whitening, normalization, etc. Autocorrelation gives a measure of matching the signal to itself while shifting the signal, known as lag, in past and future. Therefore, the autocorrelation is a function of lag. Whitening temporally decorrelates the value of signal at various instant of time thereby increasing the statistical degree of freedom. Normalization involves scaling the time and amplitude axis of the signal with respect to a reference value. Frequency domain processing may include filtering, coherence, etc. Filtering involves removing the unwanted frequency from the signals while coherence evaluates the frequencies in synchronization between two signals. The time-frequency domain processing consists of simultaneous time-frequency representation, including Spectrogram, wavelet decomposition, etc. Spectrogram is a fixed resolution time-frequency representation while wavelet offers a choice of multi-resolution. The present chapter will describe these signal processing tools like spectrogram with specific applications to EMG.

The popularity of using tri-axial accelerometers in locomotion classification has increased in recent years for clinicians and researchers. Accelerometers directly measure acceleration and hence avoid any error of calculating acceleration by integrating velocity and displacement. Accelerometers for human movement quantification generally utilize sensors like strain gauge, piezoresistive, capacitive and piezoelectric. A tri-axial accelerometer simultaneously records the acceleration in all the directions. The signal processing tools for EMG are equally applicable to accelerometers as both of them represents the movement information; however, care must be taken for low frequency and gravity biased signal from an accelerometer while selecting a signal representation method. For example, approximate entropy and fractal analysis have been considered as important time domain signal processing tools for quantification of human movement via accelerometry. Apart from locomotion classification, accelerometry has been especially advantageous in the identification of gait events like heel strike and toe off. A particular wavelet application on acceleration signals of the foot for automatic identification of toe off in locomotion and the ramp transition is presented.

In recent years, multisensory fusion and integration has shown several benefits over single sensor data for biomedical systems. In a particular case of locomotion and transition classification, the sensor fusion approach provides better classification accuracy and robust performance. In this chapter, the EMG and accelerometer performance across different time window of a gait cycle in locomotion and transition classification is presented.

This chapter presents a brief review of approaches for the EMG and accelerometer signal processing in locomotion and transition classification. It is organized in the following sections: Locomotion and transition classification, where some popular methods in use will be briefly described; Spectrogram approach for EMG signal representation, where basic principles of spectrogram approach presented; wavelet analysis of foot acceleration for gait event detection; and finally, a comparison of acceleration and EMG across gait cycle in locomotion classification is presented with an insight on sensor fusion.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

LOCOMOTION AND TRANSITION CLASSIFICATION

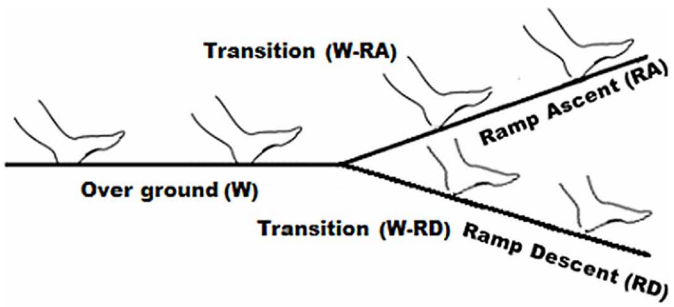
The estimated number of lower limb amputations every year in the United States is 120,000 leading to limited involvement of the individual in various physical activities (T. R. Dillingham et.al. 2002). Lower limb prostheses are designed to partially restore function in the lower limb. Passive prostheses allow for level over ground walking but do not provide net torque for walking over varied terrain. Advancements in embedded system technology have facilitated the introduction of powered prostheses which allow desired impedance for terrain walking (Au. S. et. al. 2008; F. Sup et. al. 2008). Hence, it becomes critical to accurately classify the locomotion in current powered prosthesis. Moreover, an ideal powered prosthesis must also provide the seamless transition while switching between the locomotion modes which requires to accurately identify the movement intention of amputee. While the classification problem needs to know what the amputee is currently doing; the seamless transition is an anticipation problem and therefore needs to know that what the amputee is about to do. A ramp transition is schematically shown in Figure 1.

Current powered prostheses lack in such anticipation and requires some "cueing" in the form of either manual intervention or compensatory movements which add cognitive load to the amputee resulting in non-effective use of the prosthesis. As mentioned, the seamless transition is an anticipation problem and hence becomes challenging as the future data is not available in the real time scenario. This leads to a broadened scope to investigate choice of sensors, methods in signal processing and signal representation, and various classification algorithms to solve this anticipation problem.

SENSORS

Various choices of sensors have been explored in past few decades for locomotion and transition classification. The primary criterion to select a sensor modality is the richness of information – the other criteria being the relative ease of placement (i.e. donning and doffing), the cosmesis and social acceptance (R Kobetic et. al. 2009). A detailed review of sensors for locomotion and transition classification can be found elsewhere (Michael R Tucker et. al. 2015). As EMG represents the true physiological representation of muscle contraction, it is perhaps the most promising source of information for decoding the motion intention, and potentially other application of classification. However, due to the non-stationary nature of locomotion and EMG signal being highly susceptible to noise – due to artifacts and sweating,

Figure 1. Ramp ascent and descent transition from overground walking



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

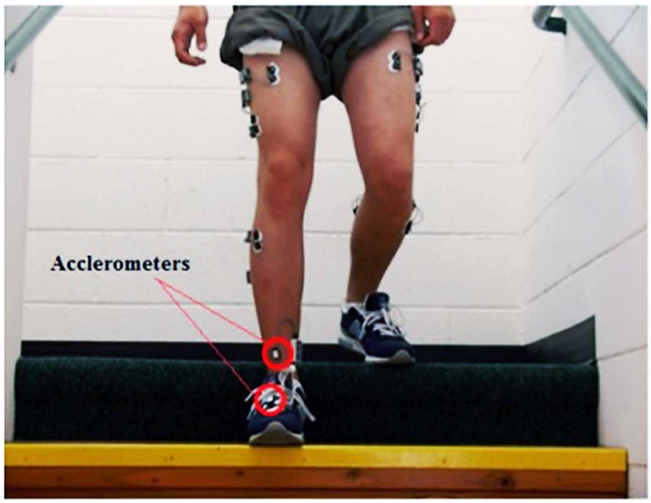
it requires extensive signal processing, followed by signal presentation, before being fed into the classifier. Particular to application in motion classification, EMG signal has shown respectable accuracy in respective applications (R. H. Chowdhary et.al. 2013; D. Graupe et. al. 1982; D. Graupe et. al. 1985; Y Huang et.al. 2005; X Zhang et. al. 2012). Due to the miniaturized sensing technology, the body mounted sensors have been used extensively in the last decade for various applications.

Specific to locomotion classification the accelerometer has been the preferred choice. A range of classifiers – including Simple threshold, Hierarchical scheme, Decision Trees, k-nearest neighborhood, Artificial Neural Network (ANN), Support Vector Machine (SVM), Naïve Bayes and Gaussian Mixture Model (GMM), Fuzzy Logic, and Markov Models (Culhane et.al. 2004; Ermes et. al. 2008; Maurer et. al. 2006; Parkka et. al. 2006; Doukas and Maglogiannis 2008; Wu et. al. 2007; Allen at. al. 2006; Salarian et. al. 2007; Pober et. al. 2006), have been used for activity classification using features from accelerometer data. A detailed review can be found elsewhere (Stephen J. Preece et.al. 2009). Subjects worn with EMG and accelerometer are shown in Figure 2.

SPECTROGRAM APPROACH FOR EMG

Due to their easy computation and reduced computation time, the time domain representation of EMG signal has been used extensively for locomotion classification (Miller et.al. 2013; Huang H et.al. 2009). However, limited signal representation in the time domain demands high number of features (Huang H. et.al. 2011), leading to high dimensionality for good classification accuracy close to 95%. Moreover, the significantly enhanced classification accuracy of activity identification by introducing frequency domain features (S J Preece et.al. 2009) is appealing for exploration of the simultaneous time-frequency representation of EMG signals in locomotion and transition classification. Such a time-frequency representation can be served by two means:

Figure 2. A subject stepping down in stair terrain wearing EMG and accelerometer sensor on foot and ankle



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. Spectrogram, and
2. Wavelet transform.

Though the wavelet transform offers multi resolution analysis in time and frequency domains compared to fixed resolution in spectrogram, wavelet transform provides redundant information in terms of a large number of wavelet coefficients, thus overburdening the classifier, and wavelet transform is suited to specific applications (Fernandes FCA et.al. 2001; Valens CA 1999). Spectrogram representation of EMG was first presented to analyze fast movement and constant contraction of human hand muscles (Hannaford B and Lehman S, 1986), however little is known about its utility in classifying locomotion modes or transitions using lower extremity EMG. A spectrogram, fixed resolution time-frequency representation, of the data segment uses periodogram approach to calculate the spectral density as following:

where N is the total number of data points in the data segment, ω is the frequency and m is the length of the window. The spectrogram is then calculated, using Matlab 2015, as the magnitude square of each periodogram as follows:

$$Spectrogram\left\{x\left(n\right)\right\}\left(m,\omega\right)=\left|P_{x}\left(m,\omega\right)\right|^{2} \tag{2}$$

The periodogram gives a good estimate of spectral density of non-stationary signals like EMG, calculating the Fourier Transform of auto-correlation to give a better estimate of spectral density. Spectrogram of a 256ms EMG with 64 ms window and 50 percent overlap is shown in Figure 3. A detailed description of periodogram can be found in (Cooper GR and McGillem CD 1999).

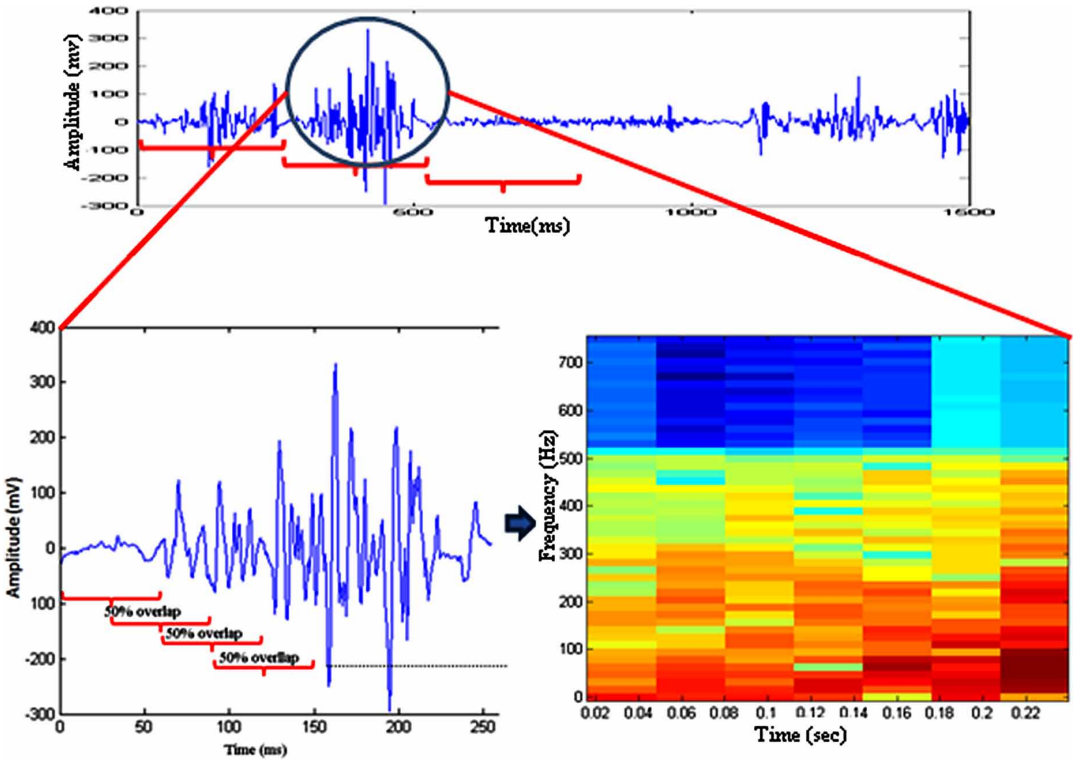
Such a periodogram is transformed to threshold based binary conversion (TBBC) by utilizing a threshold (0.5) in energy via the following equation:

$$F_{i,j}=1,\ \ if(|\ F_{i,j}\ |)^{2}\geq 0.5*\max((|\ F_{i,j}\ |)^{2})\ \ \forall i,j=0,\ \ otherwise \tag{3}$$

where F is the value of spectrogram in the bin indices, by i and j. A high threshold value of 1 may lose critical information while a lower threshold value may introduce noisy, redundant information. For this method development, the threshold value of 0.5 was selected. The spectrograms, after TBBC, from all training trials were pooled to develop the 2D histogram. Prior knowledge was extracted from the previous gait cycle. This prior knowledge was integrated into the system as follows: If the previous class is walk then there is no possibility that next class can be stair ascent. The only possibility for next class is either walk or the transition from walk to stair ascent. Hence, for a given EMG data set with prior knowledge of locomotion as walk, the classification should not return stair ascent. Thus, the prior knowledge is able to restrict the possible intended locomotion state, effectively reducing computation time and increasing classification accuracy. By training a few if-else rules using developed 2D histogram of seven lower extremities EMG, bandwidth between 3-500 Hz, namely – Tibialis Anterior (TA), Medial Gastrocnemius (MG), Rectus Femoris (RF), Vastus Lateralis (VL), Biceps Femoris (BF), Gluteus Maximus (GMax) and Gluteus Medius (GMed), of thirteen healthy subjects an overall classification error of less than 20% for over ground walking (W), stair ascent (SA) and the transition from over ground walking to stair transition (W-SA) was achieved, using leave-one-subject-out-cross validation. This error was further reduced to less than 11% by utilizing knowledge of previous gait cycle i.e. prior knowledge. More details of clas-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Spectrogram of a 256ms EMG signal with 64 ms window and 50 percent overlap



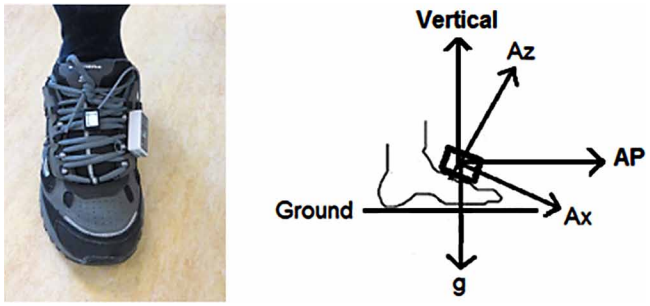
sification method and neural input can be found in (Deepak Joshi et. al. 2015). Spectrogram involves frequency transformation and hence requires obvious computational time resources. It was observed that with the computational resources of Intel Core i7 CPU @ 3.40 GHz and 8 GB RAM it took 53.5 ms and 2.5 ms for calculation of spectrogram information and classification decision, respectively. The fast classification is due to if-else rules which are computationally easy to calculate.

WAVELET APPLICATION OF ACCELEROMETRY DATA

Once the classification for locomotion is performed it is desirable to provide appropriate damping for swing and stance phase separately for the respective locomotion. Hence it becomes critical to correctly identify gait events like toe off to separate the swing and stance phase in real time. This becomes critical when the amputee is transitioning from one terrain to another. Ground reaction forces (GRF) and kinematic data from motion capture have been the golden standard for toe identification of toe off in over ground walking and treadmill walking (Y H Chang et. al. 2000; A Hreljac and R N Marshall 2000; J K De Witt 2010). Both of these sources are limited to laboratory settings and hence are not suitable for some applications like Functional Electrical Stimulation (FES) and monitoring of activity of daily living (ADL) where ambulatory monitoring systems are required. Moreover, these sources have not been tested in different terrains or in transitions. With the advent of miniaturization sensor technology and high end signal processing like wavelet it is appealing to explore inertial sensors like accelerometers

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Placement of accelerometer and associated global and local coordinate reference

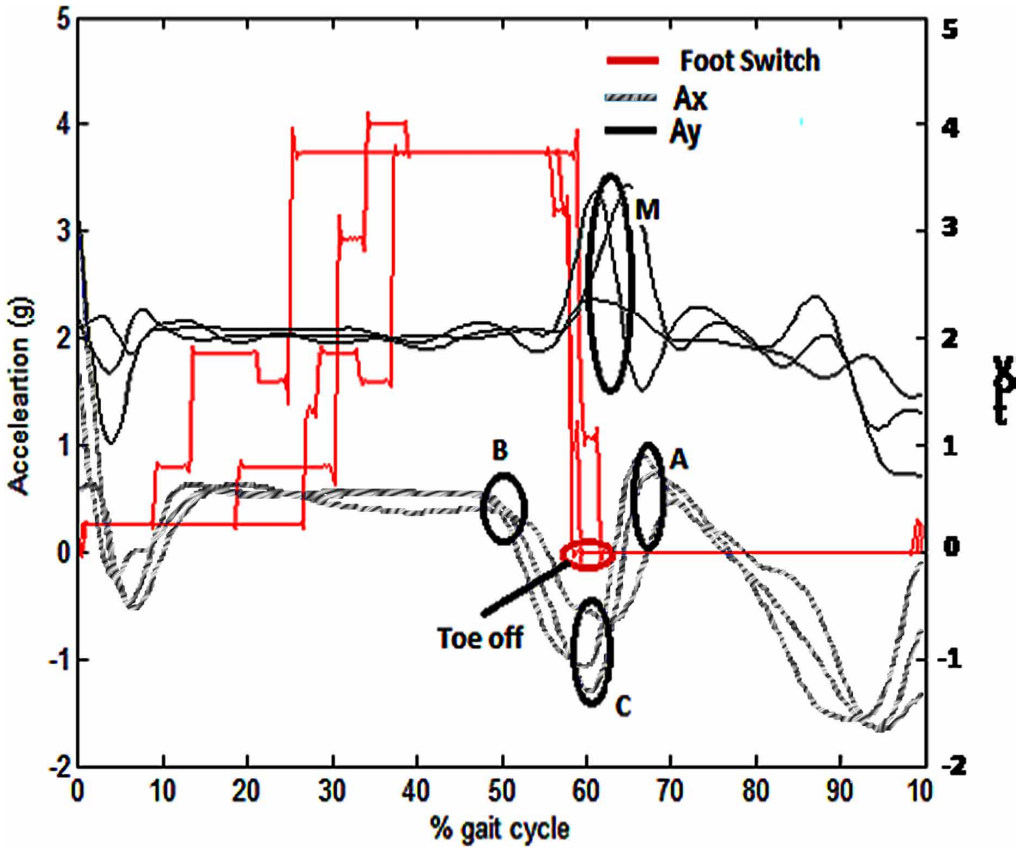


with high end signal processing for identification of toe off in different terrains and the transitions. A tri-axial accelerometer and its placement on dorsum of the foot along with its axis towards the movement are shown in Figure 4. With respect to a global coordinate reference of anterior-posterior (X) and vertical direction (Z), placement of the accelerometer on the foot led to primary measurement directions as Ax and Az (Figure 4), with Ay representing the medial-lateral direction, orthogonal to Ax and Az.

As noted in Figure 4, Az and Ax were not purely aligned with the global vertical and Z, being dependent upon the orientation of the dorsum of the foot. Foot movement during walking is generally in a low frequency range. To reduce data redundancy, we down-sampled the original data by a factor of 5, resulting in an effective sampling frequency of 300Hz. Down sampling reduced the redundancy while maintaining the critical information. The accelerometer data were filtered with a low pass second order Butterworth filter with 10Hz cut off frequency. This cutoff frequency, compared to other options like 20 and 40 Hz, avoided the local minima and maxima in the acceleration data while maintaining meaningful information of source signal to develop the heuristic algorithms. To remove the gravity component, the residual acceleration value in all directions during a period of quiet stance, is removed from the filtered data. Most of the movement of leg during locomotion is in the direction of progression i.e. Anterior Posterior (AP) and vertical. Hence, these two directional data were explored for identification of toe off event. Various signal processing algorithms and classification schemes have been employed for the identification of toe off in a gait cycle. Heuristic approaches have been especially attractive and appealing due to easy computation and ease in real time implementation algorithms (J. M. Jasiewicz et. al. 2006; A. M. Sabatini et.al. 2005). However, due to the advancement in signal processing algorithms and machine learning, researchers have been able to achieve better accuracy in toe off identification (Aung et.al. 2013). Algorithms need to be robust in identifying the toe off event in locomotion as well as in the transition gait cycle. As toe off marks the transition from stance to swing phase, a discontinuity at this transition in the vertical component of acceleration is expected. Such short period transitions are better highlighted in wavelet analysis due to the high time resolution capability for high frequency in wavelet. Further to reduce the computation time, the transition can be localized in the gait cycle using a heuristic approach before applying the wavelet transform. Figure 5 shows the toe off event localized by points A, B, C, and M. These points can be heuristically identified using the acceleration data in Ax and Az directions. Point A and B were identified utilizing mutual information in the Ax and Ay acceleration data. Specifically, point M is the maximum Ay acceleration peak in a gait cycle, and is the closest unique point in proximity to point A, which is the peak Ax acceleration. Further, point C is the maximum negative Ax peak prior to point A. We segmented Az acceleration of a gait cycle between point A and point B to locally search for the toe off.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The Ax and Ay with foot switch data during ramp walking (ascending, RA). Three separate cycles are shown from one example subject; representing those with minimum, maximum and median stride duration. The Y-axis unit for acceleration and foot switch are ±g and Volt, respectively and are shown with respect to time-normalized gait cycle. The marker C corresponds to the region of Toe-off for all the three stride cycles showing that it lies between point A and B exclusively.
Note: Acceleration in Ay direction is shown with an offset of +2g for visualization purposes.



Further, wavelet analysis can highlight the toe off event in the localized segment identified by the heuristic approach. Such an approach has been utilized recently and was demonstrated to be superior to other approaches in the past while detecting the toe off in walk (W), ramp ascent (RA), ramp descent (RD) and both the transitions (W-RA and W-RD) (Deepak Joshi et.al. 2014). Estimation error was first defined as the time difference between the toe off detection by foot switch and toe off detection by the proposed algorithm. Absolute value of this error value was calculated and averaged across the subjects to report mean absolute estimation error. Note: As the effective sampling frequency was 300 Hz, the error of a single sample of time corresponds to 3.3 ms. A comparison of accuracy in correct toe off identification is given in Table 1.

Among the other findings, the average stride durations for steady state locomotion were 1.08 ± 0.078, 1.07 ± 0.10 and 1.06 ± 0.067 seconds for W, RA and RD, respectively and 1.07 ± 0.081 and 1.06 ± 0.076 seconds for the transition W-RA and W-RD respectively. No significant differences were found between any possible pairs of within or between transition and steady state locomotion (paired t-test, α

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Comparison of estimation error (mean ± std) using the reported approaches*

Author	Sensor Placement	Over Ground (ms)	Ramp (ms)	Transition (ms)	Algorithm Type
Jasiewicz et al (2006)	Foot accelerometer	19 ±34	N/A	N/A	Heuristic
Aung et al (2013)	Foot acceleration (simulated + measured)	18.3± 27.0	5.1 ±12.1 (smooth) 15.2± 41.7 (tactile)	N/A	Heuristic + Wavelet + Machine learning
Aung et al (2013)	Ankle accelerometer	19.5± 27.1	32.5± 63.0 (smooth) 21.1± 53.1 (tactile)	N/A	Heuristic + Wavelet + Machine learning
Sabatini et al (2005)	Foot gyroscope	35	35	N/A	Heuristic
Joshi et al (2014)	Foot accelerometer	17.4 ±12.5	13.8 ±8.5	20.1 ±15.5	Heuristic + Wavelet

*The results are shown for db1 wavelet which showed better accuracy than db2 wavelet.

= 0.05). Point A, B, C and M were detected with 100 percent accuracy utilizing the Ax and Ay acceleration data; no trial was missed in examining toe off due to misidentification of any instance of point A, B or C. Hence, the Az acceleration was segmented correctly in all the trials across all subjects. The mean absolute estimation error using db2 was 25.6 ± 18.3, 32.5 ± 15.3, 33.2 ± 21.7, 32.5 ± 19.5 and 26.8 ± 16.2ms compared to 17.4 ± 12.5, 13.8 ± 8.5, 22.0 ± 16.4, 20.1 ± 15.5 and 17.1 ± 13.7ms using db1 wavelet, for W, RA, RD, W-RA, and W-RD, respectively. The db1 wavelet performed significantly better than db2 in identifying toe off with high accuracy ($p < 0.01$).

EMG AND ACCELEROMETER COMPARISION

It was concluded in a previous section that the accelerometer and EMG performed with reasonable classification accuracy in different applications for lower limb prosthetics. It will be interesting to see the performance of both sensors in a common application of classification in order to derive a fair comparison of both the sensors. Such an application is to classify the locomotion transition in grouped transition of terrain and direction. For terrain type classification, the data from W-RA and W-RD were compiled and further class labeled as ramp. Similarly, the data from W-SA and W-SD were compiled and further class labeled as stairs. For direction type classification, data from W-RA and W-SA were compiled and further class labeled as ascent. Lastly, the features from W-RD and W-SD were compiled and further class labeled as descent. In summary, the terrain group represent whether the transition is to stair or to the ramp irrespective of the direction (i.e. up and down). In a similar way the direction group represents whether the transition is to up direction or the down direction. Such a classification scheme is often described as hierarchical classification. Other than classification accuracy, it is important to see the comparison across different events of gait cycle specifically with reference to toe off as the toe off marks the transition from swing to stance phase. The toe off was marked by using foot switch data. Based on the toe off event three windows from a transitional gait cycle were extracted as follows:

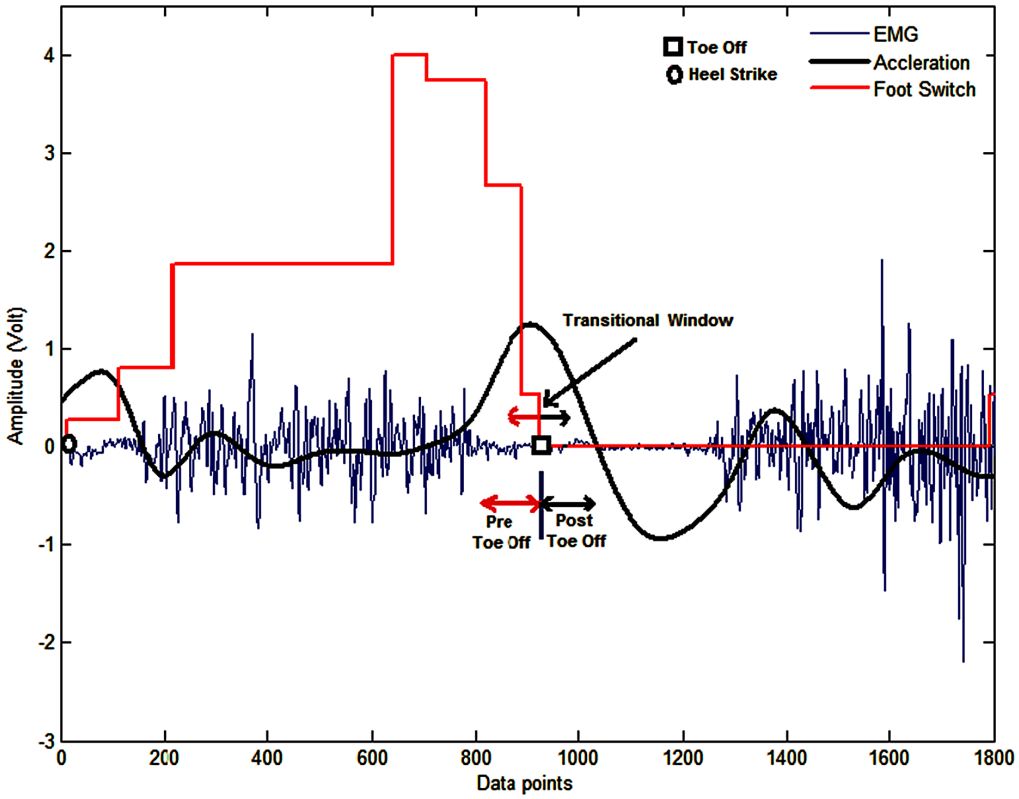
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. **Pre Toe Off:** A data segment starting from toe off and extending backward to foot strike of the transitional gait cycle. Window sizes varied from size of 50–350 ms with an increment of 50 ms.
2. **Post Toe Off:** A data segment starting from toe off and extending forward to foot strike on new terrain. Window sizes were similar to pre toe off.
3. **Transitional Window:** A combination of pre and post toe off segments.

EMG and accelerometer sensors were placed exactly as mentioned in previous sections. Data were sampled at 1500 Hz using the Noraxon Telemyo DTS system (Noraxon, Scottsdale, AZ). All post-processing was conducted using custom-written programs in Matlab (Mathworks, Natick, MA). EMG signals were band pass filtered between 20 and 500 Hz using a 4th order digital Butterworth filter. Accelerometer signals were low pass filtered with 10 Hz cut off using a 4th order digital Butterworth filter, in addition to removal of mean before further processing. Figure 6 shows the data from accelerometer and EMG sensor across the gait cycle for mentioned events.

To derive a fair comparison between both sensors it is important to represent the information from both sensors in an equal number of dimensions. Principal component analysis (PCA) is a multivariate technique that transforms the original feature space into a set of new orthogonal variables called principal components such that the variance in new dimensions is maximized. PCA finds these dimensions by directly observing the Eigen vectors of covariance matrix. For the features given in table II for both the sensors, PCA revealed that among top five Eigen vector directions EMG data could capture the 97.45%

Figure 6. EMG and accelerometer data across different events in a gait cycle



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Set of features for EMG and accelerometer

EMG (7 Channels)	Accelerometer (6 Channels)
Mean Absolute Value (MAV)	Mean
Number of Zero Crossing (ZC)	Variance
Waveform Length (WL)	Mean Trend
Number of Slope Sign Change (SSC)	Windowed Mean Difference (WMD)
Root Means Square (RMS) value	Variance Trend
Third-order Auto regressive (AR) coefficient	Windowed Variance Difference (WVD)
----------	Maximum Difference Acceleration (MDA)
Feature Space = 7 Channels X 8(5+3AR coefficient) features = 56 Dimension	Feature Space = 6 Channels X 7 features = 42 Dimension

of variance compared to just 90.36% of variance in accelerometer data (Deepak Joshi et. al. 2015). EMG data were collected from seven channels as mentioned in previous sections; however accelerometer for this study was collected from two tri-axial accelerometers on foot and shank thus constituting six channels as shown in Table 2.

Support vector Machine (SVM) and linear discrimant classifier (LDA) classifiers were used for pattern recognition. LDA is a supervised dimension-reduction technique and well-assessed classification method. The LDA involves finding a linear combination of the independent variables that minimizes the probability of misclassifying the individuals into their respective groups. It projects the dimension into one direction, useful for data classification. The ultimate goal is to find the projection to a line such that the samples from different class are well separated. SVM is a classification method that uses kernel functions to map the original features into high-dimensional feature space for constructing an optimal hyperplane to separate binary classes. A linear kernel was used for the SVM classifier; $k(x_i,x_j) = x_i^T x_j$, where x_i and x_j are the feature vectors of i^{th} and j^{th} classes, respectively, and T denotes transpose. The leave-one-out-cross validation (LOOCV) method was used to test and train the classifiers. Out of total ''n'' trials from all subjects, one trial was used for testing and n-1 trials were used for training the classifier. This procedure was iterated until all the trials had been included in training and testing. Classification accuracy (CA) was then calculated as a percentage; (number of correctly classified trials/total number of trials)*100. Comparison of classification accuracy between terrain and direction type was evaluated statistically, using a paired two-tailed *t* test. For the grouped classification (Terrain vs. Direction), overall the accelerometer performed better than the EMG in the transition gait cycle. However, the performance of both sensors was complementary to each other. In the time windows where accelerometer failed to perform well the EMG could classify the transitions with respectable accuracy. Overall, both the sensors tended toward higher accuracy for terrain type than direction type. Among the three windows the transitional window showed superior performance than the pre-toe off and post-toe off windows. A detailed analysis of the performance of both sensors in different windows of gait cycle can be found in (Deepak Joshi et. al. 2015). In this study, the toe off event during the transitional gait cycle was considered as the gold standard for separation of locomotion modes. It appears to be a good approximation for small distances, where toe off occurs close to the physical transition interface. However, if toe off occurs at a greater distance, this approximation would likely need to be modified. For a short distance, it can be

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

fairly assumed that immediately after toe off the coming locomotion mode is engaged. However, in circumstances where the distance is greater, even after toe off a major part of the gait cycle can still be considered to be within the previous locomotion mode and may violate the assumption. The potential effect of greater distances in transition should be addressed in future studies.

DISCUSSION AND CONCLUSION

The signal processing methods like spectrogram and wavelet have been presented with application to locomotion and transition classification. Frequency domain representation, spectrogram, proved to be successful in classifying the locomotion and the transition. However, it has to be yet tested for an extended number of locomotion and classification. In that case it is advised to explore a greater number of relevant features from the frequency domain representation other than the spectrogram energy only. Accelerometer and EMG data both show reasonable accuracy and complimentary information in a gait cycle. This confirms the result of a recent study (A J Young et. al. 2014) which reported that combining accelerometers and EMG data gives better classification accuracy than using either of them alone while classifying the locomotion, walk, stairs, and ramp from the data of transfemoral amputees. However, the analytical combination was not tested on locomotion transitions.

The wavelet decomposition of just one set of accelerometer data on the foot was able to achieve superior accuracy in toe off identification compared to other algorithms utilizing a machine learning approach. This study allowed us to observe that the combination of signal processing and a heuristic approach is better than a heuristic approach alone at identifying toe off using accelerometry data. This may be due to the fact that heuristic approaches are generally subjective, while signal processing approaches are data driven and hence more adaptive to signal source. The 'gold standard' for toe off detection in this study relied on foot switch data, which itself has some inherent error, due to variations in the size and shape of subjects' feet. This can lead to large variation in estimation error across subjects. It may be necessary to determine any biasing or offset between the estimations derived from foot switch and GRF plates. This offset could be further added or subtracted from the estimation of the proposed algorithm to test it against the GRF true estimation.

It was observed that throughout the transitional gait cycle the high classification performances of EMG and accelerometry were complementary to each other. This difference in window-dependent performance can be explained by the fact that EMG appears prior to the corresponding motion while acceleration is actuated during the motion, demonstrating the well-known electromechanical delay observed in skeletal muscle activation. Seemingly, an optimal combination of both sensor types across the transitional gait cycle can strengthen the revelation of transition patterns. It appears that transitions are better classified as terrain than direction type. A terrain focused approach may thus be ideal for an efficient classification structure. Our findings demonstrate that EMG and accelerometry data sources complement each other across transitional gait cycles. This presents the opportunity for an optimal sensor fusion for robust classification in application. Overall, the findings of this study suggest that a terrain and direction focused classification approach will be useful for inclusion in classification approaches utilized in lower limb amputee samples and possibly in future efforts to enable smooth locomotion transitions using powered prostheses and assistive devices.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

LIMITATIONS AND FUTURE DIRECTIONS

Able bodied participants were selected for the studies described in this chapter. With amputation the movement pattern and hence EMG and accelerometer will change as well. In future, it will be crucial to see how these sensors and the classification schemes perform with amputee data. With the complementary nature of both sensors across the gait cycle, it will be worthwhile to explore the fusion of both sensors in order to develop an optimal classification strategy. To recognize the intent of type of transition (direction/terrain) it may be important to include the visual cue as a precursor for the classification. Such a preparatory cue for obstacles and locomotion has been reported previously (R. M. Wilkie et al. 2010; R. M. Wilkie et. al. 2008), however there is a need to effectively integrate such cues into a seamless real-time classification strategy.

ACKNOWLEDGMENT

The studies described in this chapter were carried out under financial support from the U.S. Department of Defense; grant W81XWH-09-2-0144. The authors are grateful to Bryson Nakamura, Eileen Deming, and Daniel Jones for their help in data collection and post processing.

REFERENCES

Aaron, J. Y., Kuiken, T. A., & Hargrove, L. J. (2014). Analysis of using EMG and mechanical sensors to enhance intent recognition in powered lower limb prosthesis. *Journal of Neural Engineering*, *11*, 1–12.

Allen, F. R., Ambikairajah, E., Lovell, N. H., & Celler, B. G. (2006). Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models. *Physiological Measurement*, *27*(10), 935–951. doi:10.1088/0967-3334/27/10/001 PMID:16951454

Au, S., Berniker, M., & Herr, H. (2008). Powered ankle-foot prosthesis to assist level-ground and stair-descent gaits. *Neural Networks*, *21*(4), 654–666. doi:10.1016/j.neunet.2008.03.006 PMID:18499394

Aung, M. S., Thies, S. B., Kenney, L. P., Howard, D., Selles, R. W., Findlow, A. H., & Goulermas, J. Y. (2013). Automated detection of instantaneous gait events using time frequency analysis and manifold embedding. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*(6), 908–916. doi:10.1109/TNSRE.2013.2239313 PMID:23322764

Chang, Y. H., Huang, H. W., Hamerski, C. M., & Kram, R. (2000). The independent effects of gravity and inertia on running mechanics. *The Journal of Experimental Biology*, *203*, 229–238. PMID:10607533

Chowdhury, R. H., Reaz, M. B., Ali, M. A., Bakar, A. A., Chellappan, K., & Chang, T. G. (2013). Surface electromyography signal processing and classification techniques. *Sensors (Basel, Switzerland)*, *13*(9), 12431–12466. doi:10.3390130912431 PMID:24048337

Cooper, G. R., & Mc Gillem, C. D. (1999). Probabilistic methods of signal and system analysis (3rd ed.). New York: Oxford University Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Culhane, K. M., Lyons, G. M., Hilton, D., Grace, P. A., & Lyons, D. (2004). Long-term mobility monitoring of older adults using accelerometers in a clinical environment. *Clinical Rehabilitation*, *18*(3), 335–343. doi:10.1191/0269215504cr734oa PMID:15137565

De Witt, J. K. (2010). Determination of toe-off event time during treadmill locomotion using kinematic data. *Journal of Biomechanics*, *43*(15), 3067–3069. doi:10.1016/j.jbiomech.2010.07.009 PMID:20801452

Dillingham, T. R., Pezzin, L. E., & MacKenzie, E. J. (2002). Limb amputation and limb deficiency: Epidemiology and recent trends in the United States. *Southern Medical Journal*, *95*, 875–883. PMID:12190225

Doukas, C., & Maglogiannis, I. (2008). *Advanced patient or elder fall detection based on movement and sound data*. Tampere, Finland: IEEE Pervasive Health Conference and Workshops. doi:10.1109/PCTHEALTH.2008.4571042

Drost, G., Stegeman, D. F., van Engelen, B. G. M., & Zwarts, M. J. (2006). Clinical applications of high-density surface EMG: A systematic review. *Journal of Electromyography and Kinesiology*, *16*(6), 586–602. doi:10.1016/j.jelekin.2006.09.005 PMID:17085302

Ermes, M., Parkka, J., Mantyjarvi, J., & Korhonen, I. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine*, *12*(1), 20–26. doi:10.1109/TITB.2007.899496 PMID:18270033

Fernandes, F. C. A., Spaendonck, R. L. V., & Burrus, C. S. (2001). A directional, shift-insensitive, low-redundancy, wavelet transform. In *IEEE Proceedings international conference on image processing*. 10.1109/ICIP.2001.959121

Goen & Tiwari, D. C. (2013). Review of Surface Electromyogram Signals: Its Analysis and Applications. *International Journal of Electrical, Computer, Energetic*. *Electronic and Communication Engineering.*, *7*(11), 11–19.

Graupe, D., Salahi, J., & Kohn, K. H. (1982). Multifunctional prosthesis andorthosis control via microcomputer identification of temporal pattern differences in single-site myo-electric signals. *Journal of Biomedical Engineering*, *4*(1), 17–22. doi:10.1016/0141-5425(82)90021-8 PMID:7078136

Graupe, D., Salahi, J., & Zhang, D. S. (1985). Stochastic analysis of myoelectric temporal signatures for multifunctional single-site activation of prostheses and orthoses. *Journal of Biomedical Engineering*, *7*(1), 18–29. doi:10.1016/0141-5425(85)90004-4 PMID:3982004

Hannaford, B., & Lehman, S. (1986). Short time Fourier analysis of the electromyogram: Fast movements and constant contraction. *IEEE Transactions on Bio-Medical Engineering*, *33*(12), 1173–1181. doi:10.1109/TBME.1986.325697 PMID:3817851

Hreljac, A., & Marshall, R. N. (2000). Algorithms to determine event timing during normal walking using kinematic data. *Journal of Biomechanics*, *33*(6), 783–786. doi:10.1016/S0021-9290(00)00014-2 PMID:10808002

Huang, H., Kuiken, T. A., & Lipschutz, R. D. (2009). A strategy for identifying locomotion mode using surface electromyography. *IEEE Transactions on Bio-Medical Engineering*, *56*(1), 65–73. doi:10.1109/TBME.2008.2003293 PMID:19224720

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Huang, H., Zhang, F., Hargrove, L. J., Dou, Z., Rogers, D. R., & Englehart, K. B. (2011). Continuous locomotion-mode identification for prosthetic legs based on neuromuscular-mechanical fusion. *IEEE Transactions on Bio-Medical Engineering*, *58*(10), 2867–2875. doi:10.1109/TBME.2011.2161671 PMID:21768042

Huang, Y., Englehart, K. B., Hudgins, B., & Chan, A. D. (2005). A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses. *IEEE Transactions on Bio-Medical Engineering*, *52*(11), 1801–1811. doi:10.1109/TBME.2005.856295 PMID:16285383

Jasiewicz, J. M., Allum, J. H., Middleton, J. W., Barriskill, A., Condie, P., Purcell, B., & Li, R. C. T. (2006). Gait event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured individuals. *Gait & Posture*, *24*(4), 502–509. doi:10.1016/j.gaitpost.2005.12.017 PMID:16500102

Joshi, Nakamura, & Hahn. (2014). A novel approach for toe off estimation during locomotion and transitions on ramps and level ground. *IEEE Journal of Biomedical and health informatics*. Doi:10.1109/JBHI.2014.2377749

Joshi, Nakamura, & Hahn. (2015). High energy spectrogram with integrated prior knowledge for EMG-based locomotion classification. *Journal of Medical Engineering and Physics*. Doi:10.1016/j.medengphy.2015.03.001

Joshi & Hahn. (2015). Terrain and Direction Classification of Locomotion Transitions Using Neuromuscular and Mechanical Input. *Annals of Biomedical Engineering*. doi:10.100710439-015-1407-3

Kobetic, R., To, C., Schnellenberger, J., Audu, M., Bules, T., & Gaudio, R. (2009). Development of hybrid orthosis for standing, walking, and stair climbing after spinal cord injury. *Journal of Rehabilitation Research and Development*, *46*(3), 447–462. doi:10.1682/JRRD.2008.07.0087 PMID:19675995

Maurer, U., Rowe, A., Smailagic, A., & Siewiorek, D. (2006). Location and activity recognition using eWatch: a wearable sensor platform. Lecture Notes in Computer Science, 3864, 86–102.

Miller, J. D., Beazer, M. S., & Hahn, M. E. (2013). Myoelectric walking mode classification for transtibial amputees. *IEEE Transactions on Bio-Medical Engineering*, *60*(10), 2745–2750. doi:10.1109/TBME.2013.2264466 PMID:23708765

Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, *10*(1), 119–128. doi:10.1109/TITB.2005.856863 PMID:16445257

Pober, D. M., Staudenmayer, J., Raphael, C., & Freedson, P. S. (2006). Development of novel techniques to classify physical activity mode using accelerometers. *Medicine and Science in Sports and Exercise*, *38*(9), 1626–1634. doi:10.1249/01.mss.0000227542.43669.45 PMID:16960524

Sabatini, A. M., Martelloni, C., Scapellato, S., & Cavallo, F. (2005). Assessment of walking features from foot inertial sensing. *IEEE Transactions on Bio-Medical Engineering*, *52*(3), 486–494. doi:10.1109/TBME.2004.840727 PMID:15759579

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Salarian, A., Russmann, H., Vingerhoets, F. J. G., Burkhard, P. R., & Aminian, K. (2007). Ambulatory monitoring of physical activities in patients with Parkinson's disease. *IEEE Transactions on Bio-Medical Engineering*, *54*(12), 2296–2299. doi:10.1109/TBME.2007.896591 PMID:18075046

Stephen, P. (2009). Activity identification using body-mounted sensors-a review of classification techniques. *Journal of Physiological Measurement*, *30*(4), R1–R33. doi:10.1088/0967-3334/30/4/R01 PMID:19342767

Sup, F., Bohara, A., & Goldfrab, M. (2008). Design and control of a powered transfemoral prosthesis. *The International Journal of Robotics Research*, *27*(2), 263–273. doi:10.1177/0278364907084588 PMID:19898683

Tucker, M. R., Olivier, J., Pagel, A., Bleuler, H., Bouri, M., Lambercy, O., ... Gassert, R. (2015). Control strategies for active lower extremity prosthetics and orthotics: A review. *Journal of Neuroengineering and Rehabilitation*, *12*(1), 1–29. doi:10.1186/1743-0003-12-1 PMID:25557982

Valens, C.A., & Valens C. (1999). *Really friendly guide to wavelets*. Academic Press.

Wilkie, R. M., Kountouriotis, G. K., Merat, N., & Wann, J. P. (2010). Using vision to control locomotion: Looking where you want to go. *Experimental Brain Research*, *204*(4), 539–547. doi:10.100700221-010-2321-4 PMID:20556368

Wilkie, R. M., Wann, J. P., & Robert, S. (2008). Active Gaze, Visual Look-Ahead and locomotor control. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(5), 1150–1164. doi:10.1037/0096-1523.34.5.1150 PMID:18823202

Wu, W. H., Bui, A. A., Batalin, M. A., Liu, D., & Kaiser, W. J. (2007). Incremental diagnosis method for intelligent wearablesensor systems. *IEEE Transactions on Information Technology in Biomedicine*, *11*(5), 553–562. doi:10.1109/TITB.2007.897579 PMID:17912972

Zhang, X., Liu, Y., Zhang, F., Ren, J., Sun, Y. L., & Yang, Q. (2012). On design and implementation of neural-machine interface for artificial legs. *IEEE Trans IndInf*, *8*(2), 418–429.

This research was previously published in Computational Tools and Techniques for Biomedical Signal Processing edited by Butta Singh; pages 195-211; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 42
A Hybrid Kernel Extreme Learning Machine and Improved Cat Swarm Optimization for Microarray Medical Data Classification

S. Chakravarty
Orissa Engineering College, India

R. Bisoi
Siksha 'O' Anusandhan University, India

P. K. Dash
Siksha 'O' Anusandhan University, India

ABSTRACT

This paper presents the pattern classification of the binary microarray gene expression based medical data using extreme learning machine (ELM) and its variants like on-line sequential ELM (OSELM) and kernel based extreme learning machine (KELM). In the KELM category two variants namely the wavelet based kernel (WKELM) extreme learning machine and radial basis kernel extreme learning machine (RKELM) along with support vector machine (SVMRBF) and support vector machine polynomial (SVMPoly) are used to classify microarray medical datasets. Further to reduce the high dimensionality of Microarray medical datasets giving rise to high number of gene expression and small sample sizes, a modified evolutionary cat swarm optimization (MCSO) technique is adopted. The efficiency of the proposed algorithm is verified using a set of performance metrics for four binary medical datasets belonging to breast cancer, prostate cancer, colon tumor, and leukemia, respectively.

DOI: 10.4018/978-1-7998-1204-3.ch042

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

Microarray analysis and classification play very important role for early diagnosis and treatment of disease like cancer. It shows the highest rate of morbidity and mortality in both economically developed and developing countries (Kanovas, 2006) where two hundred types of cancers in human beings have been recorded using microarray technology (Ferley et al., 2012). Molecular analysis of four common types of cancers belonging to lungs, female breast, bowel and prostate (Siegel et al., 2015) reveals that their gene expression profiles and high-density DNA microarray measures (Li et al., 2011; Lee & Lee, 2003) are different and therefore they can be used to diagnose them with significant accuracy. Early detection of these cancers can prolong human life and therefore it is pertinent to classify and recognize these types of cancers mentioned above as early as possible (Li et al., 2011).

Microarray technology that produces big datasets with gene expression values for thousands of genes (6000~60000) in a cell mixture (Guyon & Andre, 2006) are economically unviable where the number of samples (n) is much larger than the number of features (p) (Bolon-canedo et al., 2014). Thus for reducing the dimensionality problem efficient feature extraction and optimal feature selection of the medical datasets are required. To overcome this problem, microarray medical datasets need dimension reduction (Bolon-Canedo et al., 2014; Saeys et al., 2007; Tan, 2006; Liu & Hiroshi, 2007). Several feature extraction techniques like the principal component analysis (PCA) (Subasi & Ismain, 2010; Lopez et al., 2009; Karg et al., 2009), kernel principal component analysis (KPCA), linear discriminate analysis (LDA) (Subasi & Ismain, 2010; Lopez et al., 2009) and canonical correlation analysis (CCA) (Vert et al., 2002) have been used widely by several researchers. However, from the extracted features an optimal feature selection approach is essential to improve the classification accuracy and its better interpretation. Feature selection methods proposed in the literature include Fast Correlation Based Filter (FCBF) (Senliol et al., 2008), Relief algorithm (Zhang et al., 2008), Support Vector Machine recursive feature elimination technique (Duan et al., 2005), sequential forward (SFS) selection and sequential backward elimination (SBE) (Peng et al., 2005). Amongst several approaches for optimal feature selection SFS and SBE outperform all others and exhibit ease of implementation and low computational complexity. However, both SFS and SBE suffer from nesting and feature deletion problems (Dunne et al., 2002), respectively, and, therefore, evolutionary techniques are used. Various evolutionary optimization techniques for feature selection include genetic algorithm (Huang et al., 2006), simulated annealing (Lin et al., 2008), ant colony optimization (Kanan et al., 2008), particle swarm optimization (Unler & Alper, 2010; Fong et al., 2013; Lie et al, 2011), differential evolution (Khusaba et al., 2011; Al-Ani et al, 2013), bacterial foraging optimization (Jakhar et al, 2011), harmony search (Dias et al., 2012), cuckoo search (Rodrigius et al., 2013), firefly (Banati et al., 2011), bat algorithm (Nakamura et al., 2012) and cat swarm optimization (Lin et al., 2009), etc.

Dimensionality reduction helps in the classification of microarray medical datasets by improving its accuracy. The important role of medical data classifier is to provide the explanation and justification for the accurate prediction of the disease (Li et al., 2011). Many traditional classifiers like KNN (Guo et al., 2003), naïve-bayes (NB) (McCallum et al., 1998), decision tree (Vlahou, et al., 2003), random forest (Diaz-Uriarte et al., 2006), ID3 (Elias, 2003), C4.5 (Quinlan, 2014) and various neural network based classifiers like multi layer perceptron (MLP) (Chaudhuri & Bhattarya, 2000), RBFNN (Subasini et al., 2009), FLANN (Misra & Dehuri, 2007), SVM (Chu & Lipo, 2005; Cortes et al., 1995; Furey et al, 2000; Hsu et al, 2003) are found in the literature. Amongst all the classifiers, ANN and its variants are extensively used by researchers to classify medical datasets (Delen et al., 2005). The success

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of the ANN based classifier is mostly dependent on the number of hidden layers, number of nodes in each hidden layer, values of the weights between input to hidden layers, hidden to output layer and the learning algorithms. In the literature, it is generally seen that when ANN is associated with gradient descent learning algorithm the performance of the model becomes time consuming and computationally prohibitive (Fernadez-nakerro et al., 2012). Besides random choice of input weights for the gradient descent learning algorithm leads to slow convergence and local minima. Thus to avoid these limitations, Huang et al. proposed a novel learning algorithm i.e. extreme learning machine (ELM) (Huang G.B. et al, 2012) using a single hidden layer architecture. In recent years ELM and its variants have gained acceptance as a powerful tool for regression and binary and multiclass classification (Zhang, 2007; Huang G.B. et al., 2012) and outperform neural networks in gene expression and medical data classification. However, amongst all the ELM variants the kernel extreme learning machine (KELM) (Sundaram et al., 2010) is immune to the choice of random weights and biases between the input and hidden layers and provides better classification accuracy without satisfying Mercer's theorem (Huang W. et al., 2010; Bencherif et al., 2015).

A modified cat swarm optimization (MCSO) technique is proposed in this paper to select the most optimal features from microarray medical datasets and Kernel Extreme Learning Machine (WKELM and RKELM) to classify the features obtained from MCSO algorithm. From recent studies it is found that CSO based feature selection outperforms DE and PSO (Chu et al., 2007; Panda et al., 2011) based approaches. Further the MPSO algorithm is expected to improve search efficiency by getting the most optimal features from the high dimensional microarray medical dataset. The proposed feature selection methods employ k-nearest neighbor algorithm as the classifier and use five-fold cross validation technique to determine the classification accuracy.

The paper is organized as follows; section 2 describes the process model and benchmark microarray medical datasets, while section 3 deals with modified cat swarm optimization based feature selection method (MCSO). All the classifiers used in this study i.e. ELM,WKELM, RKELM and SVM are discussed in section 4. While section 5 presents performance evaluation measures, computer simulation results and discussion appear in sections 6 and 7 with concluding remarks in section 8, respectively.

2. PROCESS MODEL FOR CLASSIFICATION OF MICROARRAY DATASETS

In this section, the whole process like data normalization, feature selection and classification of binary microarray medical data is described. The steps are given in Figure 1. All the datasets are normalized using max-min normalization method as shown in equation (1).

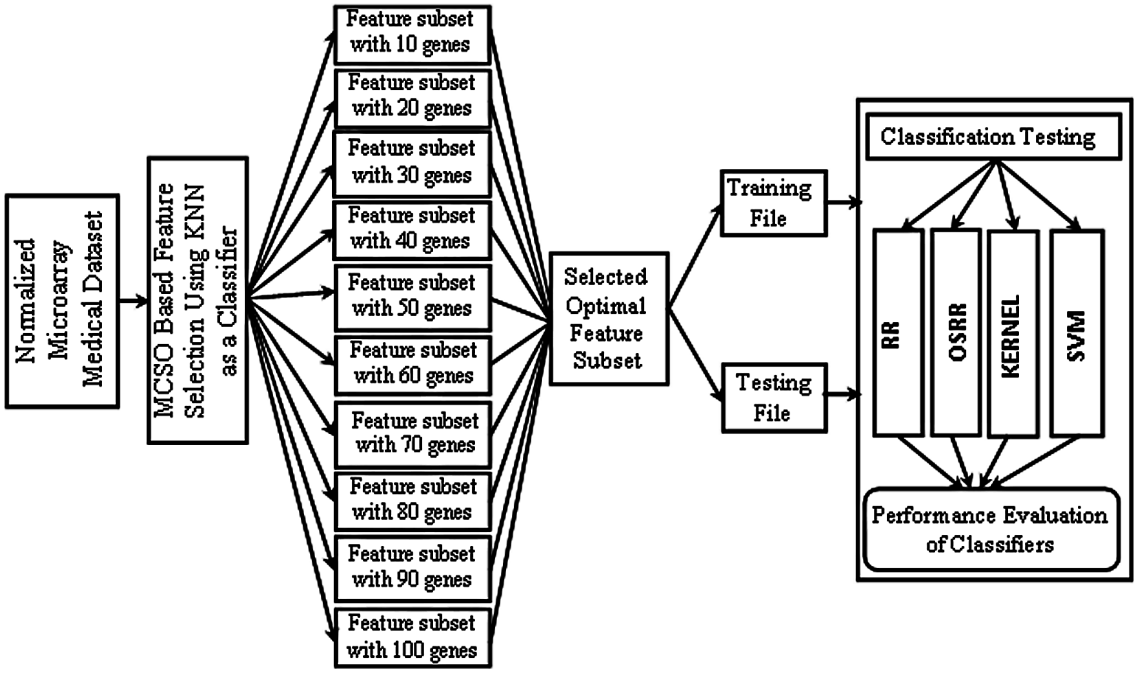
$$x_n = 2^* \frac{\left(x - x_{min}\right)}{x_{max} - x_{min}} - 1 \tag{1}$$

where x_n is the normalized form of original value x, $x_{\min}$ and $x_{\max}$ are the minimum and maximum of the dataset, respectively.

The modified cat swarm optimization algorithm (MCSO) (Saha et al.; 2013) is used to select the optimal feature subsets from these normalized datasets. For each dataset, MCSO is used to derive 10 subsets consisting of 10 to 100 genes in the interval of 10. To get the most optimal candidate features,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Process model used for classification of microarray medical datasets



k-nearest neighbor (KNN) classifier is considered to find out the classification accuracy. The subset with lower number of genes and highest classification accuracy (Statnikov et al., 2005) is chosen as the optimal candidate subset. The obtained subset is finally divided into training and testing files and used as input to all the models like ELM, WKELM, RKELM and SVM.

Four benchmark microarray binary datasets -Breast Cancer, Prostate Cancer, Colon tumor and Leukemia- are obtained from references (Li & Huiqing, 2002; Chen et al., 2012). Each dataset is divided into two data files i.e. training and testing. For binary datasets the output is either 0 or 1. Table 1 provides the features with the classes, number of samples taken for training and testing, respectively.

The Breast cancer contains 24481 features and 97 samples (Chen et al., 2012). Out of which 46 are from patients who had developed distance metastases within 5 years (labeled as "relapse"), the rest 51 samples are from patients who remained healthy after their initial diagnosis for interval of at least 5 years (labeled as "non-relapse").

Table 1. Benchmark medical datasets

Datasets	Dimension	Class	Samples	Features	Selected Features	% of Feature Reduction	Train. sample	Test. sample
Breast Cancer	97x24482	2	97	24481	50	0.2%	70	27
Prostate Cancer	136x12601	2	136	12600	50	0.4%	101	35
Colon Tumor	62x2001	2	62	2000	20	1.0%	40	22
Leukemia	76x7130	2	76	7129	100	1.4%	50	26

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The Prostate cancer data set contains 12,600 features with 136 samples in which 77 samples are prostate tumor and 59 samples are normal (Liu Bing et al., 2004). The classification task is to identify the expression patterns that correlate with the distinction of prostate tumor from the normal.

The colon cancer dataset contains 2000 genes and 62 samples collected from colon cancer patients (Liu Bing et al., 2004). Amongst them, 40 tumor biopsies are from tumors and 22 normal biopsies are from healthy parts of the colons of the same patients. Originally this dataset contains 6500 genes. Based on the confidence in the measured expression levels two thousand genes are selectedand put into the feature selection process.

The Leukemia dataset consists of 7129 genes from 47 samples of Acute Lymphoblastic Leukemia (ALL) and 25 samples of Acute Myloid Leukemia (AML) (Liu Bing et al., 2004). Figure 1 shows the process model used for classification of microarray binary medical datasets.

3. CAT SWARM OPTIMIZATION BASED FEATURE SELECTION

Feature selection is an important aspect in micro array medical data classification. The main advantage of feature selection is getting the most optimal candidate features that help in improving classification accuracy and reducing computational overhead, resource demand and storage space requirement. In the process, the most influential features get selected so that the user can be able to interpret the relation between the features and classes (Bolon-Canedo et al., 2014; Fong et al., 2013). In this study, evolutionary algorithm CSO and MCSO (Lin K.C. et al., 2014) has been applied to select the most optimal features from the binary microarray medical data sets.

The CSO (Lin K.C. et al., 2014) is a relatively new population-based, meta-heuristic evolutionary algorithm that models the behavior of cats into two modes namely the seeking mode and the tracing mode. One important characteristic of cats is that they spend most of their time in inertia and preserve their energy for future chase. In normal time their movement also very slow. When they sense a prey they chase it very quickly spending large amount of energy. The CSO refers to these two behaviors; resting with slow movement and chasing with high speed and energy as seeking and tracing modes. Similar to particle swarm optimization Cats are used in CSO as particles in D dimensions with velocities in search space. Further every cat in the CSO is associated with a fitness function and a flag to indicate whether the cat is in seeking mode or in the tracing mode. The soaring speed in seeking mode is mathematically mapped as a large change in its position.

3.1. Seeking Mode

The seeking mode has five operators namely: Seeking Memory Pool (SMP), Seeking Range of selected Dimension (SRD), Counts of Dimension to Change (CDC), Self Position Consideration (SPC) and Mixture Ratio (MR).

SMP describes the pool size of seeking memory or number of copies of a cat produced i.e. if the value of SMP is fixed as 10, than each cat is competent to store 10 solution sets as candidates.

SRD is used to state the mutative portion for the selected dimensions i.e. if a dimension is selected for mutation, the difference between the old and new values may not be out of range.

CDC corresponds to the number of dimensions to be transformed in the seeking process.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

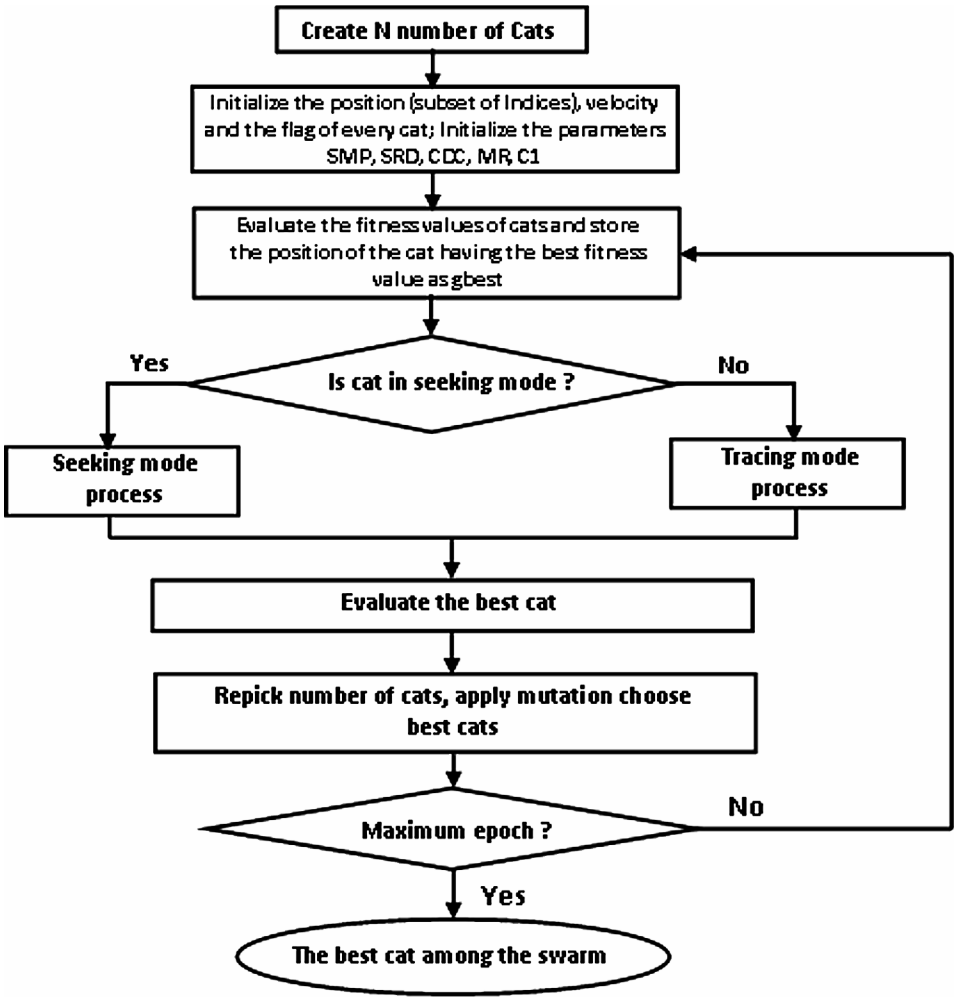
SPC is a Boolean value and if it is true, one position within the memory will save the current solution set and not be changed.

MR is a very small value i.e. fraction of population is used to guarantee that the cats spend most of their time in seeking mode.

The flow chart of CSO is given in Figure 2 and the steps executed in seeking mode are given as:

Step 1: Number of cats i.e. N are created by initializing the position, velocities and the flag of each cat.
Step 2: Pick a number of cats and according to MR put them into tracing and seeking mode.
Step 3: Evaluate the fitness value based on its positions.
Step 4: If the termination condition is met, final solution will be the best position of each cat in the solution space. Otherwise, go to Step 2.

Figure 2. Flow chart for Cat Swarm Optimization



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.2. Tracing Mode

Tracing mode is very much similar to the local search method of swarm in particle swarm optimization algorithm. In this mode, cats trace the targets with high energy by changing the positions with their own velocities. Position and velocity of the i^{th} cat in the D dimensional solution space are represented as

$$X_i=[X_{i1},X_{i2},X_{i3},.....X_{iD}] \tag{2}$$

$$V_i=[V_{i1},V_{i2},V_{i3},.....V_{iD}] \tag{3}$$

Global best position of the cat swarm is represented as

$$gbest=[gbest_1,gbest_2,.....gbest_D] \tag{4}$$

Update the velocity and position of the current cat using equations (5) and (6)

$$V_{iD} = w * V_{iD} + c_1 * r_1 * (gbest_D - X_{iD}) \tag{5}$$

$$X_{iD} = X_{iD} + V_{iD} \tag{6}$$

where w is the inertia weight, c_1 is the acceleration constant and r_1 is a random number uniformly distributed in the range (0,1). A large inertia weight facilitates a global search while a small inertia weight facilitates a local search. In this study, w is set as 0.9 initially. Normally the inertia weight has a large value in the beginning of the search and it decreases as the global best solution is achieved. Let $F_{g,k}$ is the best fitness value corresponding to $gbest_D$ and fitness differential between kth and k-1th iteration is obtained as

$$\Delta F_k = \left|F_{g,k} - F_{g,k-1}\right|, \text{ and } s=\Delta F_k/f_1,\ s\leq 1 \text{ and } s>0 \tag{7}$$

Here s signifies the convergence speed and f is obtained as

$$f = \max[\Delta F_1, \Delta F_2,\ldots,\Delta F_k] \tag{8}$$

In swarm optimization the diversity of population of cats is computed using the standard deviation of fitness values of population and is obtained as

$$\sigma = \sqrt{(1/pop)\sum_{j=1}^{pop}(\frac{F_{j,k} - F_{avg,k}}{ff})}, \quad F_{avg,k} = (1/pop)\sum_{j=1}^{pop} F_{j,k} \tag{9}$$

$$ff = \max[(F_k(1) - F_{avg,k}),(F_k(2) - F_{avg,k}),.....,(F_k(pop) - F_{avg,k})] \tag{10}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

where $F_{avg,k}$ is the average population fitness in the kth iteration. $F_k(i)$ is the ith particle fitness at the kth iteration. Finally the inertia weight is iteratively tuned as

$$w_k = w_{k-1} + \alpha s + \beta\sigma^2 \tag{11}$$

where α, β are factors which are to be chosen. Here $\sigma^2>0$,$s>0$, $0.4\leq w\leq 0.9$.

In equation (5) c1 is an acceleration coefficient to influence the velocity of the cat in the solution space and is normally chosen as 2.0. However to provide adaptive nature to CSO, this is updated as

$$c_1 = c_0 - \frac{(k_{\max} - k)}{2k_{\max}}, c_0 = 2 \tag{12}$$

A cat swarm represents a set of indices. Using these indices, corresponding reduced feature subset is derived from the original dataset. It may happen that at the time of selection one or more number of indices may not fall within the range of columns present on that particular dataset. The following formula is adopted to select every new index value within the range of columns.

if index_value≤max_*column* && *index_value*≥0.25*max_*column*

new_index =index_value-SRD% of in*dex_value* *(13)*

else

*new_index =index_va*lue+SRD% of *index_value* *(14)*

*To get most optimal cand*idate features with better classification accuracy, Modified Cat Swarm Optimization (explained in next section) is applied to all the datasets.

3.3. Modified Cat Swarm Optimization (MCSO)

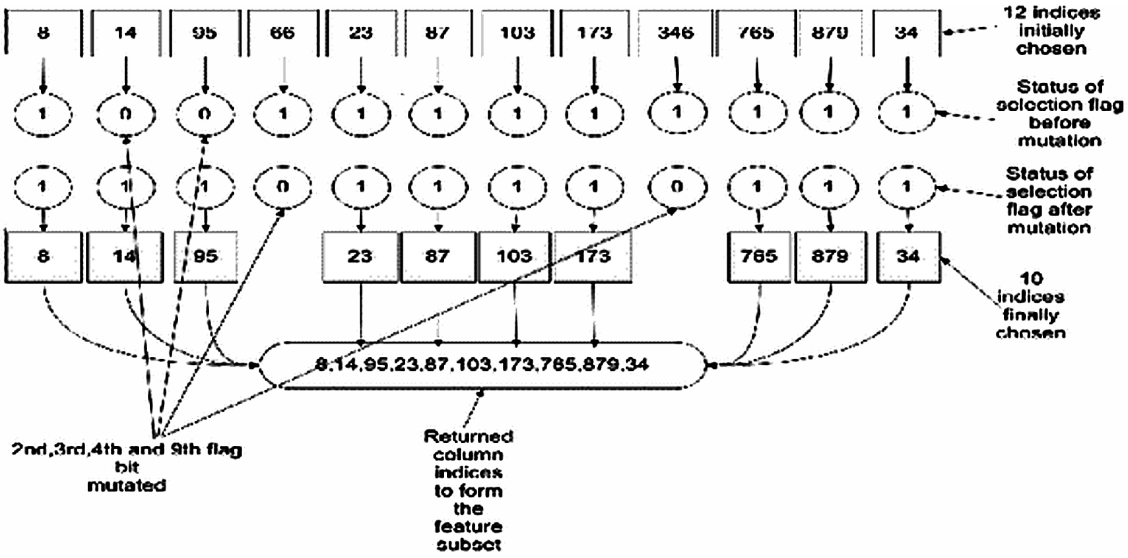
This study proposes a modified cat swarm optimization (MCSO) to improve searching ability in the locality of the best cats. In cat swarm optimization, tracing mode helps the cats in updating their velocities using equation 5. Once the best cats are chosen based on their fitness, the CSO method re-picks the number of cats and on the basis of mixture ratio (MR) distributes them into seeking mode and tracing mode. But in MCSO method, the re-picked cats are again subjected to mutation (Lin K.C. et al., 2014) and in the process randomly chosen indices are mutated. The mutation operation in MCSO is briefly described in the following sub section.

3.3.1. Mutation Operation in MCSO

Mutation operation helps in generating new solution sets in the vicinity of other solution sets. As mutation is applied randomly over the feature subset, every feature has a chance to be altered i.e. from selected to unselected and vice versa. At first, all the indices have been marked by setting their individual status

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Introduction of mutation in MCSO algorithm



flag as 1 (selected) and 0 (unselected).Initially for generating an optimal feature subset consisting of 10 features, 12 indices are chosen by MCSO in which 10 indices are marked as selected i.e. 1 and two indices are marked as unselected i.e. 0.The procedure is explained in Figure 3. When mutation is applied to 2nd,3rd indices, their stat[us] h[av]e been changed to 'selected' and status of 4th and 9thindices become u[ns]electe[d.] Finally, the mutated subsets with minimum features with maximum accuracy are selected for the real classification using ELM, WKELM, RKELM and SVM.

While determining the fitness of the cats, K-Nearest Neighbor (KNN) (Nakamura et al., 2012), the most popular and traditional classifier, has been chosen to determine classification accuracy of the feature subset. In KNN, 5-fold cross validation technique is used to determine the classification accuracy.

3.3.2. Modified Cat Swarm Optimization Algorithm

```
Begin
       Generate an Initial Swarm of 'n' no. of cats or different solutions xi
(i=1, 2,..,n)
       Initialize  position(subset of indices), velocity,  flag of every cat
       Initialize  parameters- SMP,SRD,CDC,MR,C1
       while (it<NI)
              do
                     distribute the cats into seeking/tracing mode
                            while (t< Max_Cats)
                                   do
                                          Get a Catt and Evaluate its fitness
Fi using KNN
                                          if  (Catt is in seeking mode) then
                                          Search  by seeking mode process
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
else
Search  by tracing mode process
                              end if
                              Update the best cats/solutions
                            Repick the number of cats and apply mutation
                     find the current best cat among the swarm
end while
       end while
    Generate the optimal feature subset using the selected indices
end
```

In Table 2, number of gene selected and the classification accuracies obtained from both CSO and MCSO using KNN classifier for all four data sets are depicted. Finally, the highest classification accuracy with the minimum number of genes are selected and considered for the real classification using ELM, WKELM, RKELM and SVM models. Figure 4 shows the number of selected gene and the corresponding classification accuracy obtained from KNN algorithm and MCSO based feature selection technique.

4. ELM, KERNEL ELM AND SVM CLASSIFIERS

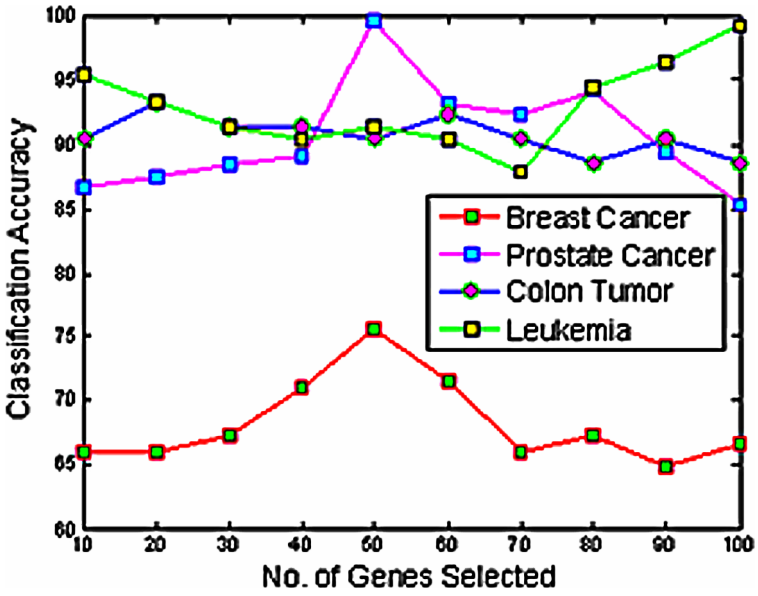
This section discusses all four classifiers -ELM, Kernel ELM and SVM- used to classify binary microarray medical datasets.

Table 2. Number of gene selected and classification accuracy using KNN from all the datasets after CSO, MCSO based Feature Selection

Genes Selected	Breast Cancer		Prostate Cancer		Colon Tumor		Leukemia	
	CSO	MCSO	CSO	MCSO	CSO	MCSO	CSO	MCSO
10	62.28	66.06	76.56	86.67	81.27	90.48	86.67	95.38
20	61.13	66.06	77.23	87.54	85.56	**93.33**	85.45	93.32
30	60.34	67.27	75.86	88.44	82.24	91.43	85.64	91.42
40	67.75	70.91	76.43	89.21	83.32	91.43	84.43	90.48
50	60.35	**75.76**	81.28	**99.56**	81.23	90.48	84.12	91.42
60	64.47	71.52	77.56	93.21	84.23	92.38	82.89	90.42
70	62.76	66.06	76.65	92.38	83.36	90.48	81.16	87.86
80	63.38	67.27	74.45	94.22	79.38	88.54	87.78	94.42
90	57.78	64.85	79.12	89.43	82.17	90.48	89.47	96.48
100	69.92	66.66	75.35	85.33	81.12	88.56	91.11	**99.23**

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. No. of gene selected and classification accuracy (MCSO)



4.1. Extreme Learning Machine

Extreme learning machine (ELM) originally proposed by G.B. Huang (Huang G.B et al., 2012) for single layered feed forward neural networks (SLFN) provides efficient learning paradigms when applied in clustering, regression and classification of real world problems. The hidden layer parameters of the ELM are computationally free from iterations that make ELM extremely fast by significantly reducing the computational time required to train the SLFNs. The SLFN model often requires high number of hidden neurons while generating optimal solutions. Further the hidden layer parameters can be chosen in a random manner independent of the training samples. Figure 5 shows the architecture of the ELM chosen for classification of binary microarray medical datasets. The output function of the ELM with L hidden nodes can be represented by the following equations.

$$f_L(x) = \sum_{i=1}^{L} \beta_i g_i x = \sum_{i=1}^{L} \beta_i G(a_i, b_i, x), x \in R^d, \beta_i \in R^m \tag{15}$$

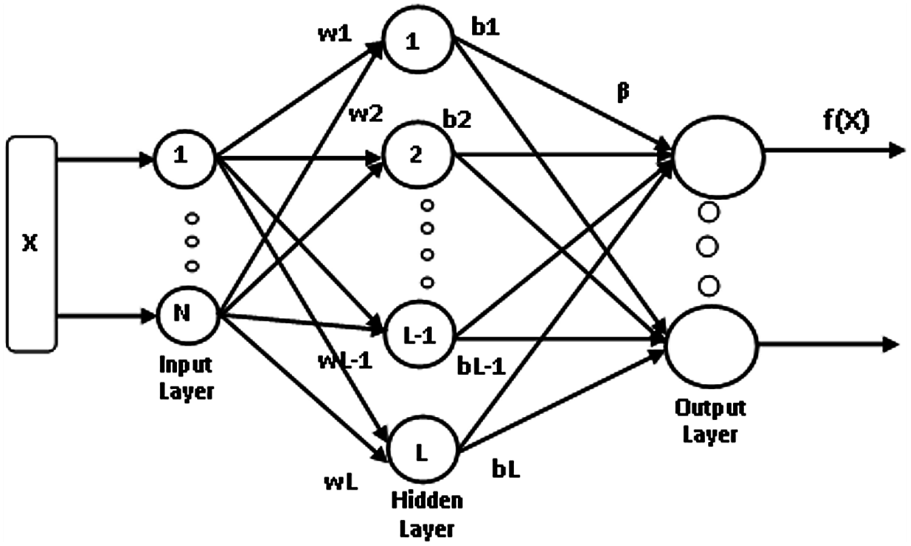
For additive nodes with activation function g, g is defined as

$$g_i = G(a_i, b_i, x) = \mathrm{g}(a_i x + b_i), a_i \in R^d, b_i \in R \tag{16}$$

$$\sum_{i=1}^{L} \beta_i G(a_i, b_i, x) = t_j, j = 1, \ldots\ldots\ldots N \tag{17}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Architecture of ELM



The above equation is rewritten as

$$H\beta = T \tag{18}$$

where

$$H(w_1 \ldots w_L, b_1 \ldots b_L, x_1 \ldots x_N) = \begin{pmatrix} g(w_1.x_1 + b_1) & \cdots & g(w_L.x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1.x_N + b_1) & \cdots & g(w_L.x_N + b_L) \end{pmatrix} \tag{19}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix} \tag{20}$$

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix} \tag{21}$$

H represents the hidden layer output matrix of the neural network, *T* is the target vector, and β represents the weight vector between the hidden layer neurons and the output layer neuron. It has been shown in reference (Huang G.B et al., 2012) that given an arbitrary small value $\in > 0$, if the hidden layer activation function is infinitely differentiable, the number of hidden neurons $L \leq N$, the input weights

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and hidden layer biases can be evaluated, then ELM can be trained to solve a linear system optimization problem using the following equation.

$$|| H \hat{\beta} - T ||= \min_{\beta} || \mathrm{H}\beta - T || \tag{22}$$

where $\hat{\beta}$ is obtained as,

$$\hat{\beta} = H^{\dagger}T = (H^{T}H)^{-1}H^{T}T \tag{23}$$

and is the smallest norm least square solution of w_i=T; $H^{\dagger}$ represents the Moore Penrose generalized inverse of H.

The procedure of ELM can be summarized in the following three steps:

Step 1: Choose the input weights w_i and hidden layer biases b_i randomly.
Step 2: Compute the hidden layer output matrix H.
Step 3: Obtain the output weight $\hat{\beta}$ using the equation $\hat{\beta} = H^{\dagger}T$

Here the function $g(w_1x_1+b_1)$ in equation is chosen as

$$g(w_1x_1+b_1) = \tan h(w_1x_1+b_1) \tag{24}$$

and w_1 and x_1 are the first weight and pattern vectors, respectively.

4.2. Kernel Extreme Learning Machine

In standard ELM, the role of the hidden layer is to map the input layer neurons to the hidden layer neurons i.e. hidden layer neurons of the ELM are mapping data from the data space to some higher dimensional space, where each dimension corresponds to a hidden neuron. Hence, the performance of ELM mostly depends upon the hidden layer neurons and it is application specific. To avoid the above said hidden neuron selection problem a newly developed kernel ELM (Lin Bin et al., 2014) (KELM) has been used to classify all microarray medical datasets considered in this study. In KELM, a positive regularization coefficient 'C' is introduced to make it more generalized and stable. Now the output weight β is

$$\beta = H^{T}\left(\frac{1}{C} + HH^{T}\right)^{-1} T \tag{25}$$

here H is the hidden layer output matrix, C is the regulation coefficient and T is the output matrix. The output function of ELM becomes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$f(x) = h(x)H^T \left(\frac{1}{C} + HH^T\right)^{-1} T \tag{26}$$

Here, instead of knowing the hidden layer feature mapping, $h(x)$, its corresponding $k(u,v)$ can be computed. The absence of L i.e. hidden layer neurons in KELM further simplifies Kernel ELM computation process. The kernel matrix of ELM based on Mercer's condition can be defined as

$$\theta_{ELM} = HH^T : \theta_{ELMi,j} = h(x_i).h(x_j) = k(x_i, x_j) \tag{27}$$

Hence the output of the kernel based ELM is written as

$$f(x) = h(x)H^T \left(\frac{1}{C} + HH^T\right)^{-1} T = \begin{bmatrix} k(x, x_1) \\ \bullet \\ \bullet \\ k(x, x_N) \end{bmatrix}^T \left(\frac{1}{C} + \theta_{ELM}\right)^{-1} T \tag{28}$$

where $\theta_{ELM}=HH^T$ and $k(x_i, x_j)$ is the kernel function of hidden neurons of single hidden layer feed forward neural network. Amongst the different kernel functions satisfying Mercer's conditions available in the literature the wavelet kernel (WKELM) and radial basis function kernel (RKELM) are considered in this study. Radial basis kernel function is a typical local kernel function in which λ and Υ are used as the parameters. On the other hand, the complex wavelet kernel function uses vector i.e. [d, e, f] as parameters. Depending on the datasets, right choice of kernel function and proper tuning of the parameters are highly required to get better results.

- Radial basis kernel

$$k(x, y) = e^{\left(-a\|x-y\|\right)} \tag{29}$$

- Wavelet kernel

$$k(x, y) = \cos\left(d\frac{\|x-y\|}{e}\right) e^{\left(-\frac{\|x-y\|^2}{f}\right)} \tag{30}$$

Kernel ELM is advantageous in comparison to ELM because there is no requirement of knowing the hidden layer feature mapping and setting the number of hidden layer neurons L. It achieves better generalization, more stability compared to ELM and is faster than SVM.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. PERFORMANCE EVALUATION METRICS

Performance of all the classifiers are evaluated by different measures like training accuracy, testing accuracy, confusion matrix, Receiver Operating Characteristic curve (ROC), sensitivity, specificity, Gmean, and F-score, etc.

5.1. Confusion Matrix

Confusion matrix represents the values of TP, TN, FP, FN (Abdulahamit, 2004; Ferri et al., 2009; Bowes et al., 2012). TP (True Positive) - correctly classified positive samples, TN (True Negative) - correctly classified negative samples, FP (False positive) - misclassified negative samples and FN (False negative) - misclassified positive samples.

5.2. Classification Accuracy

The classification Accuracy (Abdulahamit, 2004; Malathi et al., 2010) is a commonly used evaluation criterion for a standard classification system.

$$\text{Accuracy}= \frac{TP+TN}{TP+TN+FP+FN} \tag{31}$$

5.3. Sensitivity

Sensitivity (Abdulahamit, 2004) is also known as TPR (true positive rate) or recall. Sensitivity evaluates how well a model can recognize abnormal records. Higher the sensitivity is better the classifier becomes. It always identifies proportion of accurately classified samples to total samples.

$$\text{Sensitivity}= \frac{TP}{TP+FN} \tag{32}$$

5.4. Specificity

Specificity is regarded as the TNR (true negative rate). Specificity measures the ability of the proposed method to identify normal cases. It identifies accurately classified negative samples to total samples. If specificity becomes high like sensitivity and their difference is very less (<1%), then the classifier is a better one. Consider

$$\text{Specificity}= \frac{TN}{TN+FP} \tag{33}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5.5. Gmean

Gmean (Bowes et al., 2012) is used when the combined performance of the two classes need to be judged by a single metric. It measures both the sensitivity and specificity in a balanced manner.

$$Gmean = \sqrt{Sensitivity * Specificity} \tag{34}$$

5.6. F-Score

Like Gmean, F-score (Ferri et al., 2009; Bowes et al., 2012) is also an important metric to check the combined performance.

F-score =(2*Sensitivity*Specificity)/(Sensitivity+Specificity) (35)

5.7. ROC

ROC curve (Slaby, 2007) is a graphical representation of sensitivity against (1-specificity).

6. SIMULATION RESULTS

In this section, obtained results from all six models -ELM, OSELM, WKELM, RKELM, SVMRBF and SVMPoly–and the datasets considered in this study have been given. All the classifiers are implemented in the following environment, Operating System: Windows XP professional, CPU: Intel Core i3-370M (2.4 GHz), and Memory: 4GB RAM.

The number of hidden nodes 'L' in ELM is varying from 2 to 22 with an increment of 2 neurons each time. A vector consisting of three parameters [p1, p2, p3] is given as an input to WKELM. The parameters p1 and p3 are kept constant as 1 and 10, respectively whereas p2 is varied from 4 to 44 at an increment of 4. In RKELM, λ is varied as 2^1, 2^1.5,……,2^6 and ϒ is varied as 2^-1, 2^-1.5,……2^-6, respectively. Simulation results for binary class datasets in respect of ELM, OSELM, kernel extreme learning machines like WKELM and RKELM are demonstrated through graphs and Tables. Figures 6-9 show testing accuracy of ELM, OSELM, WKELM, and RKELM, respectively. Since from Figure 7 it is observed that OSELM does not perform as well as the ELM, subsequent performance measures like Sensitivity and specificity of only ELM, WKELM and RKELM for all binary class datasets are shown in Figures 10-15, respectively. Figure 16 represents sensitivity vs specificity vs Gmean of WKELM of Breast cancer dataset. Figures 17 and 18 represent ROC for WKELM and RKELM of all the datasets, respectively. Results obtained from all the datasets like time required for training (Trg time), training accuracy (Trg Acc), testing accuracy (Tst Acc) and standard deviation of testing accuracy (stdv Tst Acc), TP, FN, FP, TN, sensitivity, specificity, Gmean, F-Score and Wt-norm, training and testing accuracy for ELM is demonstrated in Tables 3 to 6 along with the training and testing accuracy of OSELM, respectively. Tables 7-10 and 11-14 display results obtained from WKELM, and RKELM, respectively.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

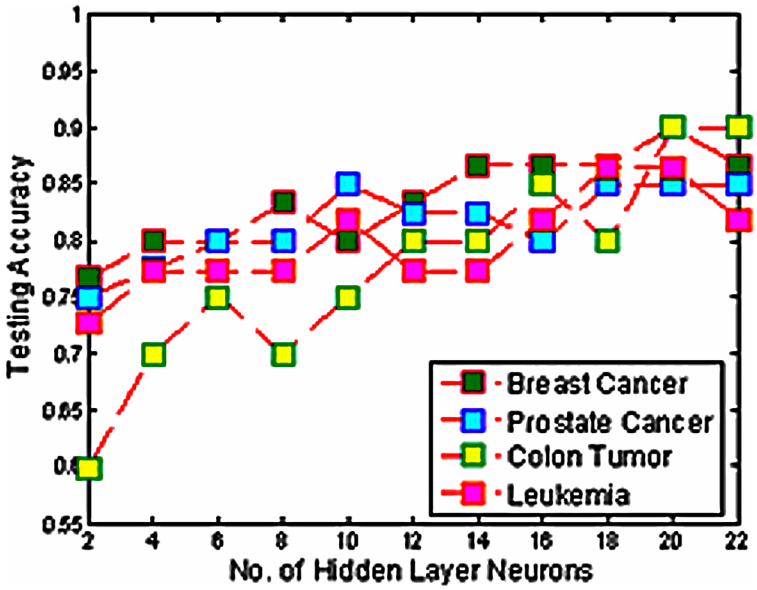
All microarray medical datasets are classified using SVM scikit learn (Pedregosa et al., 2011). Two types of kernels i.e. RBF and Polynomial are chosen. The simulation results like precision, recall, F_1-score, classification accuracy, AUC values and confusion matrices for binary appear in Table 15.

7. DISCUSSION

In this paper, extreme learning machine variants like ELM, OSELM, WKELM, RKELM, along with support vector machines models like SVMRBF and SVMPoly are used to classify binary microarray medical datasets very effectively. The evolutionary meta-heuristic algorithm like CSO and MCSO are used to select the most optimal candidate subsets. Number of gene selected and classification accuracy obtained from KNN in feature selection step are analyzed and compared. From Table 2, it is clearly demonstrated that MCSO yields better classification results in comparison with CSO.

Maximum classification accuracy and AUC for all the models and datasets are given in Table 16. It is observed from the Tables that the Kernel ELM exhibits testing accuracy of 97, 97, 95, 95 percent for breast cancer, prostate cancer, colon tumor, leukemia respectively. This clearly demonstrates the superiority of KELM in comparison with ELM, OSELM, SVMRBF and SVMPoly models. Further amongst the two variants of KELM i.e. WKELM and RKELM, the former produces better accuracy than that of the latter in all datasets except one i.e. prostate cancer. Maximum classification accuracy obtained from all the models and datasets are represented through a bar graph shown in Figure 19. Also it is seen from Figures 10-15 WKELM produces higher sensitivity and specificity as compared to other models. Also WKELM shows better ROC, in comparison to that of RKELM as demonstrated in Figures 17 and 18. KELM gives more AUC value in all the four datasets of breast cancer, prostate cancer, colon tumor, leukemia considered in this study as compared to those belonging to ELM, OSELM,SVMRBF and SVMPoly, respectively. Further, it is seen from the literature that the classification accuracies ob-

Figure 6. Testing Accuracy (ELM)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Testing Accuracy (OSELM)

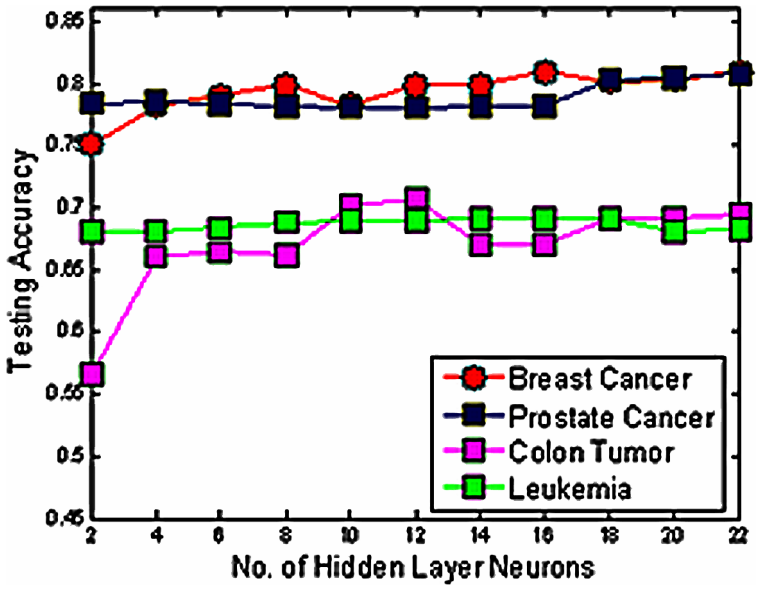
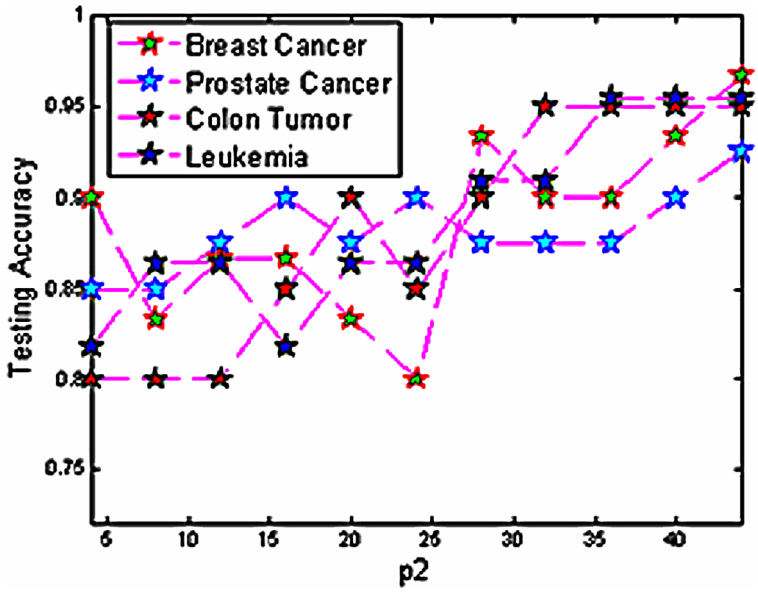


Figure 8. Testing Accuracy (WKELM)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Testing Accuracy (RKELM)

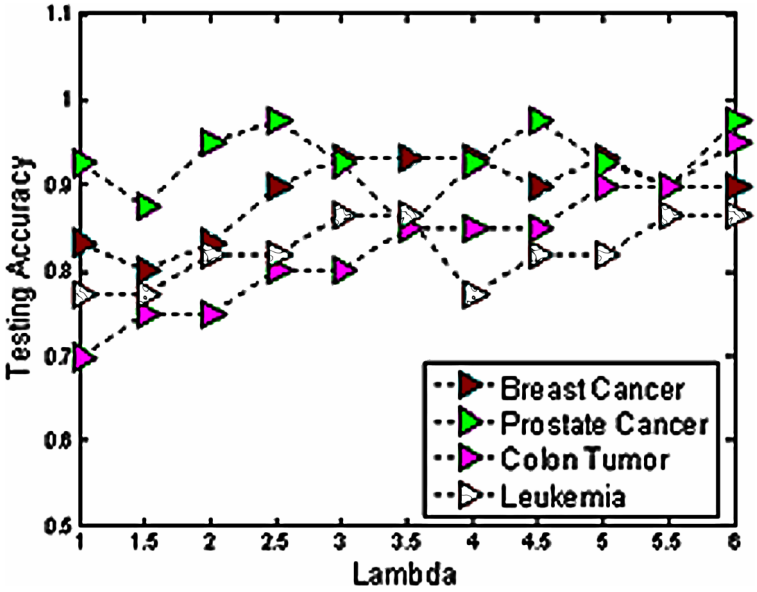
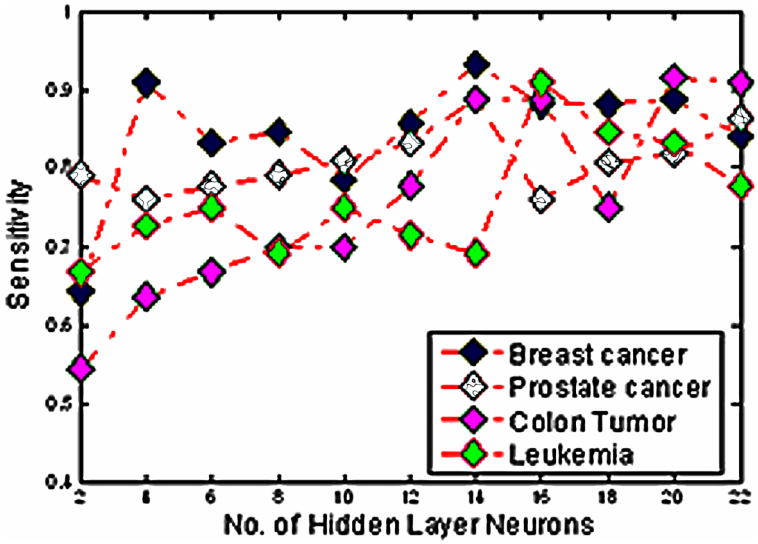


Figure 10. Sensitivity (ELM)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Sensitivity of (WKELM)

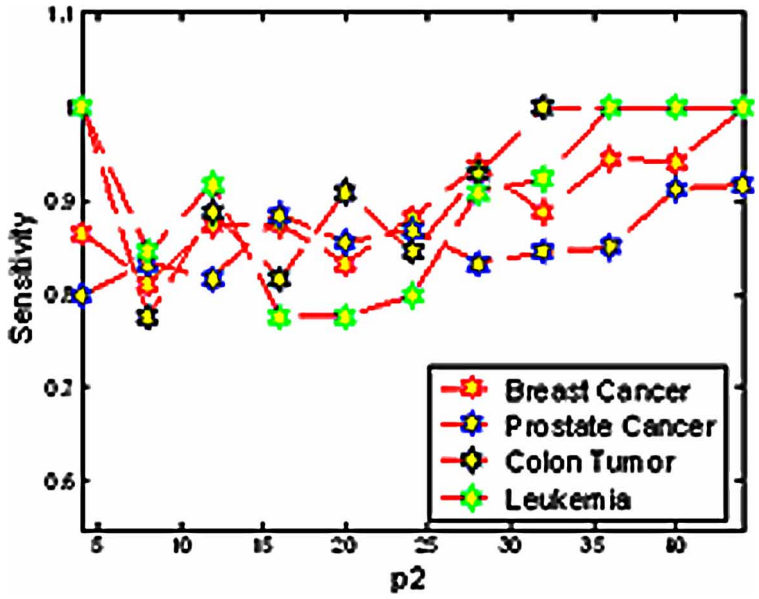
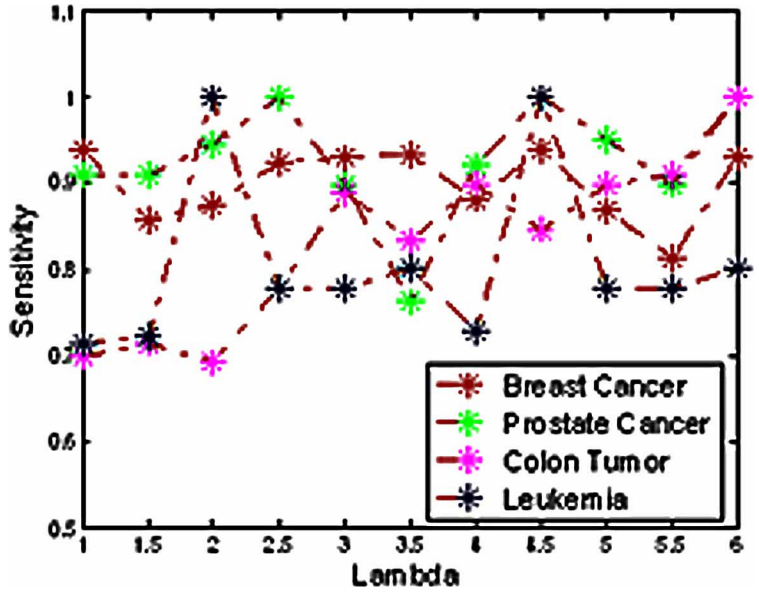


Figure 12. Sensitivity (RKELM)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 13. Specificity (ELM)

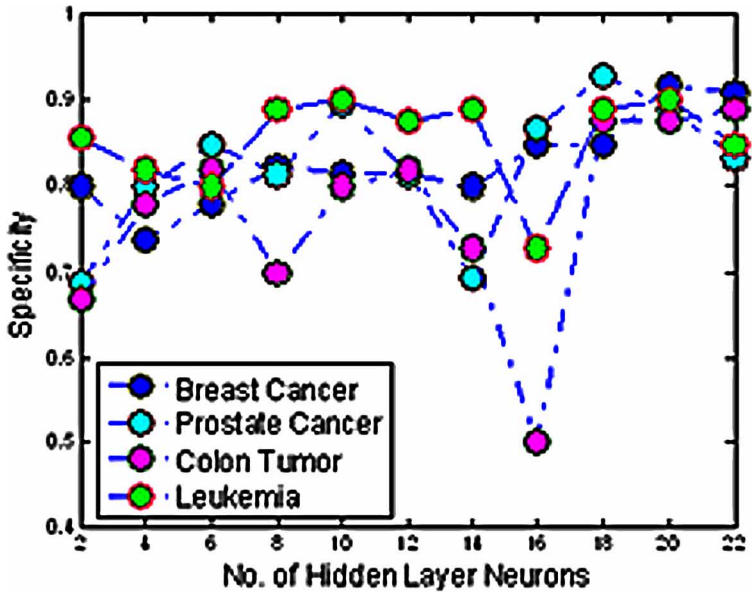
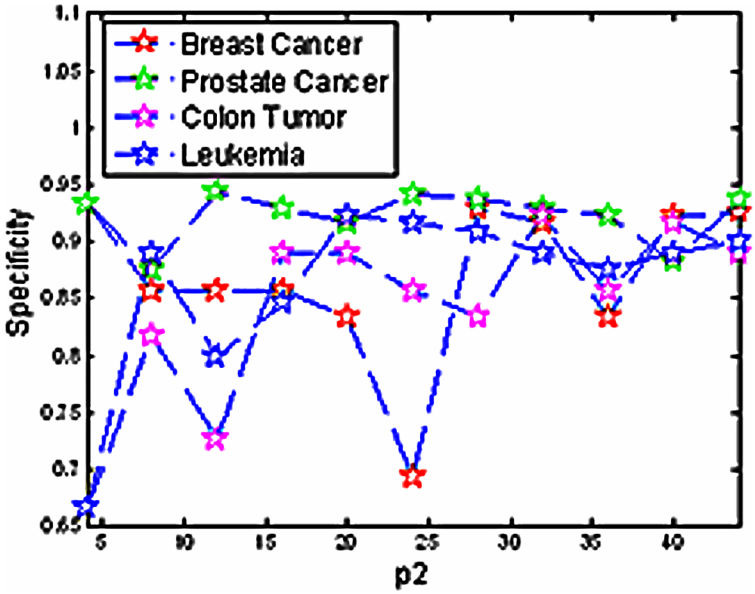


Figure 14. Specificity (WKELM)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. Specificity (RKELM)

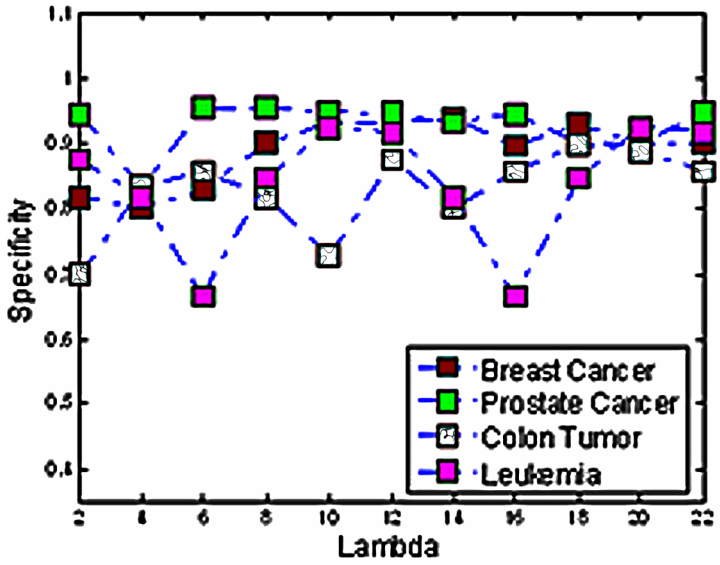
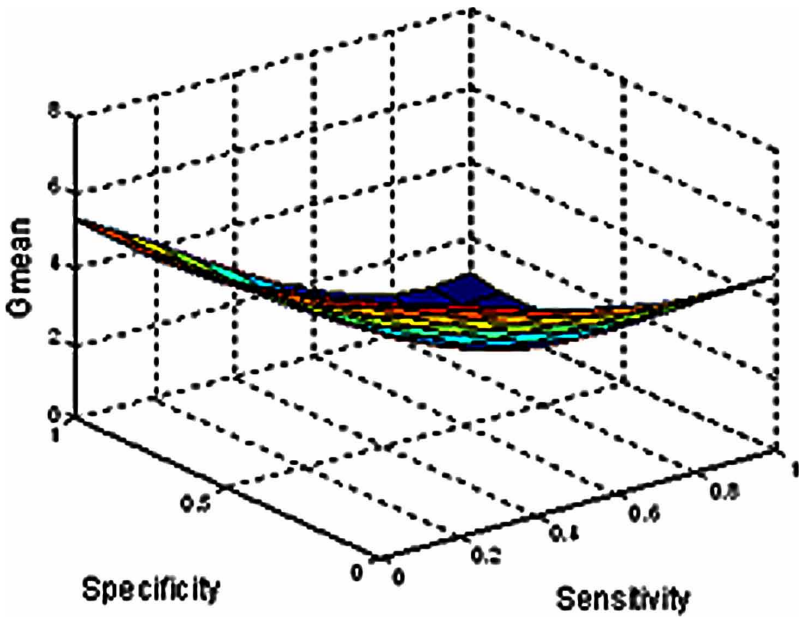


Figure 16. Sensitivity vs Specificity vs Gmean, WKELM (Breast Cancer)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 17. ROC, WKELM

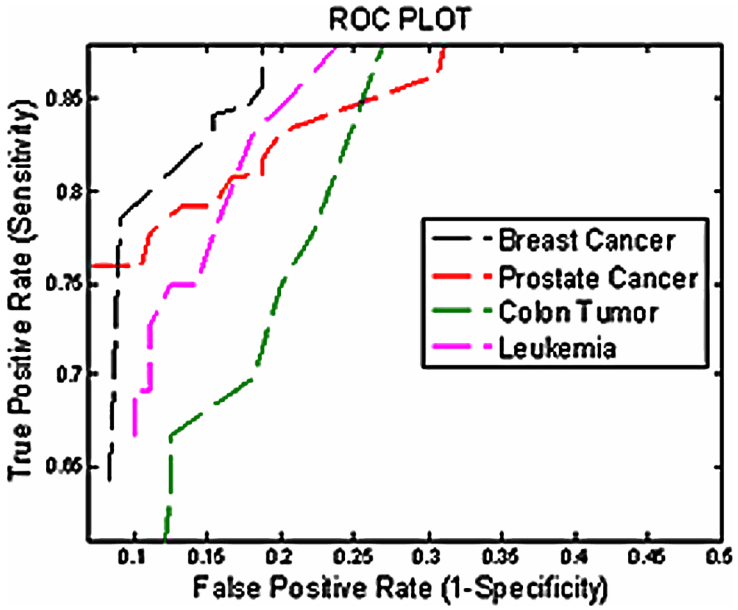
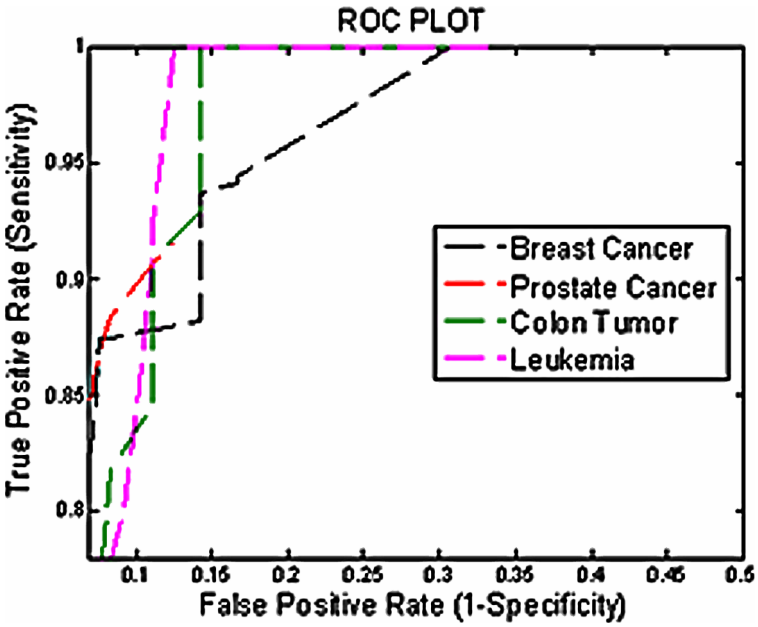


Figure 18. ROC, RKELM



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of ELM and Training Accuracy, Testing Accuracy of OSELM (Breast Cancer)

L	ELM													OSELM	
	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	Gmean	FScore	Wt-Norm	Trg Acc	Tst Acc
2	.0109	.4851	.7667	.0532	8	0	1	12	.6429	.8000	.7454	.7407	1.03e+11	.4779	.7500
4	.0109	.5111	.8000	.0376	10	1	5	14	.9091	.7368	.8184	.8140	2.2e+11	.4808	.7810
6	.0093	.5000	.8000	.0509	10	2	4	14	.8333	.7778	.8051	.8046	5.2e+10	.4882	.7900
8	.0140	.4962	.8333	.0501	11	2	3	14	.8462	.8235	.8348	.8347	2.8e+11	.4782	.7980
10	.0218	.5000	.8000	.0308	11	3	3	13	.7857	.8125	.7990	.7989	3.2e+11	.4838	.7811
12	.0062	.5592	.8333	.0491	12	2	3	13	.8571	.8125	.8345	.8342	1.09e+11	.4778	.7985
14	.0031	.5481	.8667	.0446	14	1	3	12	.9333	.8000	.8641	.8615	5.4e+11	.4556	.7980
16	.0124	.4962	.8667	.0242	15	2	2	11	.8824	.8462	.8641	.8639	3.72e+11	.4676	.8086
18	.0171	.4888	.8667	.0343	15	2	2	11	.8824	.8462	.8641	.8639	1.23e+11	.4989	.8003
20	.0031	.5481	.9000	.0361	16	2	1	11	.8889	.9167	.9027	.9026	3.2e+11	.4332	.8029
22	.0062	.5587	.8667	.0342	16	3	1	10	.8421	.9091	.8750	.8743	3.1e+11	.4431	.8088

Table 4. Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of ELM and Training Accuracy, Testing Accuracy of OSELM (Prostate Cancer)

L	ELM													OSELM	
	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	Gmean	FScore	Wt-Norm	Trg Acc	Tst Acc
2	0	.9305	.7500	.0322	19	5	5	11	.7917	.6875	.7377	.7359	11.89	.9328	.8440
4	.0031	.9611	.7750	.0357	19	6	3	12	.7600	.8000	.7797	.7795	21.20	.9444	.8526
6	.0187	.9583	.8000	.0399	21	6	2	11	.7778	.8462	.8112	.8105	16.15	.9443	.8823
8	.0031	.9666	.8000	.0387	19	5	3	13	.7917	.8125	.8020	.8019	21.34	.9423	.8812
10	.0093	.9527	.8500	.0383	17	4	2	17	.8095	.8947	.8511	.8500	28.73	.9411	.8804
12	.0187	.9583	.8250	.0381	20	4	3	13	.8333	.8125	.8229	.8228	17.41	.9389	.8805
14	.0194	.9694	.8250	.0382	24	3	4	9	.8889	.6923	.7845	.7784	26.86	.9332	.8811
16	.0194	.9694	.8000	.0411	19	6	2	13	.7600	.8667	.8116	.8098	31.51	.9298	.8815
18	.0166	.9666	.8500	.0398	21	5	1	13	.8077	.9286	.8660	.8639	27.94	.9299	.8831
20	.0361	.9361	.8500	.0395	18	4	2	16	.8182	.8889	.8528	.8521	31.60	.9338	.8743
22	.0555	.9555	.8500	.0397	19	3	3	15	.8636	.8333	.8483	.8482	18.60	.9331	.8765

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of ELM and Training Accuracy, Testing Accuracy of OSELM (Colon Tumor)

L	ELM													OSELM	
	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	Gmean	FScore	Wt-Norm	Trg Acc	Tst Acc
2	0	.6227	.6000	.2503	6	5	3	6	.5455	.6667	.6030	.6000	5.0e+3	.5887	.5656
4	.0187	.5590	.7000	.2207	7	4	2	7	.6364	.7778	.7035	.7000	5.6e+3	.5876	.6600
6	.0062	.5727	.7500	.2108	6	3	2	9	.6667	.8182	.7385	.7347	7.5e+3	.5656	.6643
8	.0031	.5772	.7000	.1897	7	3	3	7	.7000	.7000	.7000	.7000	5.7e+3	.5454	.6612
10	.0031	.5363	.7500	.2000	7	3	2	8	.7000	.8000	.7483	.7467	4.6e+3	.5854	.7008
12	.0062	.6000	.8000	.2004	7	2	2	9	.7778	.8182	.7977	.7975	3.2e+3	.5834	.7076
14	.0062	.5318	.8000	.2032	8	1	3	8	.8889	.7273	.8040	.8000	3.7e+3	.5647	.6700
16	.0078	.5227	.8500	.1998	8	1	2	9	.8889	.5000	.6124	.6000	6.4e+3	.5776	.6703
18	.0109	.5727	.8000	.1902	9	3	1	7	.7500	.8750	.8101	.8077	6.03e+3	.5665	.6905
20	.0124	.6090	.9000	.2448	11	1	1	7	.9167	.8750	.8956	.8953	3.14e+3	.5578	.6899
22	.0109	.5818	.9000	.2231	10	1	1	8	.9091	.8889	.8989	.8989	6.4e+3	.5678	.6956

Table 6. Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of ELM and Training Accuracy, Testing Accuracy of OSELM (Leukemia)

L	ELM													OSELM	
	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	Gmean	FScore	Wt_Norm	Trg Acc	Tst Acc
2	0	.8346	.7273	.0879	10	5	1	6	.6667	.8571	.7559	.7500	309.0	.7886	.6809
4	.0046	.8115	.7727	.0358	8	3	2	9	.7273	.8182	.7714	.7701	462.7	.7554	.6812
6	.0031	.8153	.7727	.0214	9	3	2	8	.7500	.8000	.7746	.7742	696.1	.7550	.6834
8	.0093	.8846	.7727	.0126	9	4	1	8	.6923	.8889	.7845	.7784	559.8	.7543	.6876
10	.0046	.8923	.8182	.0754	9	3	1	9	.7500	.9000	.8216	.8182	187.2	.7400	.6888
12	.0031	.8615	.7727	.0983	10	4	1	7	.7143	.8750	.7906	.7865	124.3	.7372	.6888
14	.0093	.8923	.7727	.0897	9	4	1	8	.6923	.8889	.7845	.7784	8.052	.7223	.6909
16	.0031	.9153	.8182	.0878	10	1	3	8	.9091	.7273	.8131	.8081	12.70	.7467	.6909
18	.0062	.9153	.8636	.0756	11	2	1	8	.8462	.8889	.8673	.8670	3.695	.7609	.6907
20	.0156	.8527	.8636	.0698	10	2	1	9	.8333	.9000	.8660	.8654	78.61	.7656	.6812
22	.0124	.9611	.8182	.0760	7	2	2	11	.7778	.8462	.8112	.8105	19.35	.7653	.6814

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. Parameters, Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of WKELM (Breast Cancer)

λ	Parameters (d,e,f)	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	T	FN	FP	TN	Sens.	Spec.	G mean	F score	Wt-Norm
1	[1,4,10]	.0020	1	.9000	.0413	13	2	1	14	.8667	.9333	.8994	.8988	6.0113
1	[1,8,10]	.0044	1	.8333	.0511	13	3	2	12	.8125	.8571	.8345	.8342	6.1120
1	[1,12,10]	.0044	1	.8667	.0915	14	2	2	12	.8750	.8571	.8660	.8660	6.1521
1	[1,16,10]	.0045	1	.8667	.0782	14	1	1	14	.8750	.8571	.8660	.8660	6.1589
1	[1,20,10]	.0046	1	.8333	.0987	15	3	2	10	.8333	.8333	.8333	.8333	6.1867
1	[1,24,10]	.0044	1	.8000	.0987	15	2	4	9	.8824	.6923	.7816	.7759	6.1560
1	[1,28,10]	.0065	1	.9333	.0928	15	1	1	13	.9375	.9286	.9330	.9330	6.1987
1	[1,32,10]	.0044	1	.9000	.1085	16	2	1	11	.8889	.9167	.9027	.9026	6.1789
1	[1,36,10]	.0044	1	.9000	.0425	17	1	2	10	.9444	.8333	.8872	.8854	6.1962
1	[1,40,10]	.0044	1	.9333	.0937	16	1	1	12	.9412	.9231	.9321	.9320	6.2091
1	[1,44,10]	.0053	1	.9667	.0988	16	0	1	13	1	.9256	.9636	.9630	6.3214

Table 8. Parameters, Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of WKELM (Prostate Cancer)

λ	Parameters	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	G mean	F score	Wt-Norm
1	[1,4,10]	.0052	1	.8500	.0161	20	5	1	14	.8000	.9333	.8641	.8615	4.089
1	[1,8,10]	.0090	1	.8500	.0161	20	4	2	14	.8333	.8750	.8539	.8537	4.092
1	[1,12,10]	.0051	1	.8750	.0161	18	4	1	17	.8182	.9444	.8790	.8768	4.070
1	[1,16,10]	.0053	1	.9000	.0131	23	3	1	13	.8846	.9286	.9063	.9061	4.142
1	[1,20,10]	.0049	1	.8750	.0246	24	4	1	11	.8571	.9167	.8864	.8859	4.152
1	[1,24,10]	.0049	1	.9000	.0322	20	3	1	16	.8696	.9412	.9047	.9040	4.220
1	[1,28,10]	.0053	1	.8750	.0246	20	4	1	15	.8333	.9375	.8839	.8824	3.967
1	[1,32,10]	.0049	1	.8750	.0246	22	4	1	13	.8462	.9286	.8864	.8854	4.065
1	[1,36,10]	.0057	1	.8750	.0263	23	4	1	12	.8519	.9231	.8867	.8860	3.950
1	[1,40,10]	.0050	1	.9000	.0322	21	2	2	15	.9130	.8824	.8976	.8974	4.103
1	[1,44,10]	.0060	1	.9250	.0324	22	2	1	15	.9167	.9375	.9270	.9270	4.008

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 9. Parameters, Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of WKELM (Colon Tumor)

λ	Parameters	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	G mean	F score	Wt-Norm
1	[1,4,10]	.0090	.6727	.8000	.1019	8	0	4	8	1	.6667	.8165	.8000	9.911
1	[1,8,10]	.0086	.6363	.8000	.2020	7	2	2	9	.7778	.8182	.7977	.7975	.9971
1	[1,12,10]	.0104	.7000	.8000	.2020	8	1	3	8	.8889	.7273	.8040	.8000	917.08
1	[1,16,10]	.0086	.7272	.8500	.2034	9	2	1	8	.8182	.8889	.8528	.8521	910.74
1	[1,20,10]	.0091	.7272	.9000	.2750	10	1	1	8	.9091	.8889	.8889	.8889	877.60
1	[1,24,10]	.0086	.7363	.8500	.4040	11	2	1	6	.8462	.8571	.8516	.8516	900.57
1	[1,28,10]	.0084	.6909	.9000	.2020	13	1	1	5	.9286	.8333	.8797	.8784	382.08
1	[1,32,10]	.0085	.7000	.9500	.1010	7	0	1	12	1	.9231	.9608	.9600	376.34
1	[1,36,10]	.0087	.6909	.9500	.2020	13	0	1	6	1	.8571	.9258	.9231	381.06
1	[1,40,10]	.0097	.6909	.9500	.2030	7	0	2	11	1	.9167	.9574	.9565	369.88
1	[1,44,10]	.0098	.6909	.9500	.2032	11	0	1	8	1	.8889	.9428	.9412	369.88

Table 10. Parameters, Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of WKELM (Leukemia)

λ	Parameters	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	G mean	F score	Wt-Norm
1	[1,4,10]	.0041	1	.8182	.0470	10	0	4	8	1	.6667	.8165	.8000	4.1028
1	[1,8,10]	.0044	1	.8636	.0941	11	2	1	8	.8462	.8889	.8673	.8670	3.9651
1	[1,12,10]	.0041	1	.8636	.0576	11	1	2	8	.9167	.8000	.8563	.8544	4.0346
1	[1,16,10]	.0041	1	.8182	.0744	7	2	2	11	.7778	.8462	.8112	.8105	4.0346
1	[1,20,10]	.0042	1	.8636	.0470	7	2	1	12	.7778	.9231	.8473	.8442	4.1711
1	[1,24,10]	.0043	1	.8636	.0880	8	2	1	11	.8000	.9167	.8563	.8544	4.1028
1	[1,28,10]	.0041	1	.9091	.0880	10	1	1	10	.9091	.9091	.9091	.9091	3.9651
1	[1,32,10]	.0045	1	.9091	.0470	12	1	1	8	.9231	.8889	.9058	.9057	3.9651
1	[1,36,10]	.0040	1	.9545	.0576	14	0	1	7	1	.8750	.9354	.9333	4.0346
1	[1,40,10]	.0042	1	.9545	.0744	13	0	1	8	1	.8889	.9428	.9412	4.1028
1	[1,44.10]	.0043	1	.9545	.0744	12	0	1	9	1	.9000	.9487	.9474	4.1008

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11. Parameters (λ, ϒ), Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of RKELM (Breast Cancer)

λ	ϒ	Trg Time	Trg Acc	TstAcc	StdvTstAcc	TP	FN	FP	TN	Sens.	Spec.	G mean	F Score	Wt-Norm
2^1	2^-1	.0019	1	.8333	.0413	10	4	1	15	.7143	.9375	.8183	.8108	8.208
2^1.5	2^-1.5	.0019	1	.8000	.0511	12	4	2	12	.7500	.8571	.8018	.8000	8.228
2^2	2^-2	.0016	1	.8333	.0753	11	3	2	14	.7857	.8750	.8292	.8290	10.95
2^2.5	2^-2.5	.0016	1	.9000	.1006	11	5	2	12	.8824	.9231	.9025	.9023	9.862
2^3	2^-3	.0017	1	.9333	.0742	15	1	1	13	.9375	.9286	.9330	.9330	10.95
2^3.5	2^-3.5	.0016	1	.9333	.0732	14	1	1	14	.9333	.9333	.9333	.9333	10.95
2^4	2^-4	.0017	1	.9333	.1011	13	0	2	15	1	.8824	.9393	.9375	11.59
2^4.5	2^-4.5	.0020	1	.9000	.0985	12	2	1	15	.8571	.9375	.8964	.8955	11.74
2^5	2^-5	.0016	1	.9333	.0985	15	0	2	13	1	.8667	.9309	.9286	10.93
2^5.5	2^-5.5	.0016	1	.9000	.1032	14	0	3	13	1	.8125	.9014	.8966	11.32
2^6	2^-6	.0019	1	.9000	.0996	14	2	1	13	.8750	.9286	.9014	.9010	11.59

Table 12. Parameters (λ, ϒ) Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of RKELM (Prostate Cancer)

λ	ϒ	Trg Time	Trg Acc	TstAcc	Stdv TstAcc	TP	FN	FP	TN	Sens.	Spec.	G mean	F Score	Wt-Norm
2^1	2^-1	.0056	1	.9250	.0161	20	2	1	17	.9091	.9444	.9266	.9264	5.832
2^1.5	2^-1.5	.0065	1	.8750	.0263	20	2	3	15	.9091	.8333	.8704	.8696	7.055
2^2	2^-2	.0056	1	.9500	.0246	17	1	1	21	.9444	.9545	.9495	.9495	7.612
2^2.5	2^-2.5	.0073	1	.9750	.0241	18	0	1	21	1	.9545	.9770	.9767	7.805
2^3	2^-3	.0060	1	.9250	.0131	18	2	1	19	.9000	.9500	.9247	.9243	7.99
2^3.5	2^-3.5	.0044	1	.8500	.0246	16	5	1	18	.7619	.9474	.8496	.8446	9.543
2^4	2^-4	.0047	1	.9250	.0131	23	7	1	14	.9200	.9333	.9266	.9266	9.381
2^4.5	2^-4.5	.0050	1	.9750	.0161	22	0	1	17	1	.9444	.9718	.9714	11.29
2^5	2^-5	.0046	1	.9250	.0263	19	1	2	18	.9500	.9000	.9247	.9243	11.74
2^5.5	2^-5.5	.0047	1	.9000	.0161	18	2	2	18	.9000	.9000	.9000	.9000	4.27
2^6	2^-6	.0048	1	.9750	.0246	21	0	1	18	1	.9474	.9733	.9730	9.481

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 13. Parameters (λ, ϒ), Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of RKELM (Colon Tumor)

λ	ϒ	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	G mean	F Score	Wt-Norm
2^1	2^-1	.0080	.6181	.7000	.2020	7	3	3	7	.7000	.7000	.7000	.7000	1.30e+3
2^1.5	2^-1.5	.0152	.7454	.7500	.1010	10	4	1	5	.7143	.8333	.7715	.7692	1.11e+3
2^2	2^-2	.0085	.6272	.7500	.1010	9	4	1	6	.6923	.8571	.7703	.7660	1.00e+4
2^2.5	2^-2.5	.0090	.8636	.8000	.1010	7	2	2	9	.7778	.8182	.7977	.7975	1.18e+5
2^3	2^-3	.0081	.8363	.8000	.1007	8	1	3	8	.8889	.7273	.8040	.8000	1.30e+5
2^3.5	2^-3.5	.0080	.8363	.8500	.2020	10	2	1	7	.8333	.8750	.8539	.8537	1.27e+5
2^4	2^-4	.0081	.8363	.8500	.1002	9	1	2	8	.9000	.8000	.8485	.8471	1.29e+5
2^4.5	2^-4.5	.0081	.8363	.8500	.1010	11	2	1	6	.8462	.8571	.8516	.8516	1.33e+5
2^5	2^-5	.0107	.8454	.9000	.1001	9	1	1	9	.9000	.9000	.9000	.9000	1.29e+5
2^5.5	2^-5.5	.0119	.8181	.9000	.1010	10	1	1	8	.9091	.8889	.8989	.8989	1.00e+5
2^6	2^-6	.0080	.8363	.9500	.2020	13	0	1	6	1	.8571	.9258	.9231	1.23e+7

Table 14. Parameters (λ, ϒ), Training Time, Training Accuracy, Testing Accuracy, Standard Deviation of Testing Accuracy, TP, FN, FP, TN, Sensitivity, Specificity, Gmean, FScore and Norm of RKELM (Leukemia)

λ	ϒ	Trg Time	Trg Acc	Tst Acc	Stdv Tst Acc	TP	FN	FP	TN	Sens.	Spec.	G mean	F Score	Wt-Norm
2^1	2^-1	.0039	1	.7727	.0576	10	4	1	7	.7143	.8750	.7906	.7865	3.3237
2^1.5	2^-1.5	.0038	1	.7727	.0470	8	3	2	9	.7213	.8182	.7714	.7701	6.4376
2^2	2^-2	.0038	1	.8182	.0576	10	0	4	8	1	.6667	.8165	.8000	6.5151
2^2.5	2^-2.5	.0039	.9824	.8182	.0526	7	2	2	11	.7778	.8462	.8112	.8105	4.80e-4
2^3	2^-3	.0039	.9789	.8636	.0600	7	2	1	12	.7778	.9231	.8473	.8442	.0048
2^3.5	2^-3.5	.0038	.9824	.8636	.0941	8	2	1	11	.8000	.9167	.8563	.8544	.0080
2^4	2^-4	.0041	.9649	.7727	.0288	8	3	2	9	.7273	.8182	.7714	.7701	.0091
2^4.5	2^-4.5	.0038	.9929	.8182	.0288	10	0	4	8	1	.6667	.8165	.8000	.0101
2^5	2^-5	.0038	.9614	.8182	.0372	7	2	2	11	.7778	.8462	.8112	.8105	.0104
2^5.5	2^-5.5	.0044	.9964	.8636	.0235	7	2	1	12	.7778	.9231	.8473	.8442	.0117
2^6	2^-6	.0039	.9719	.8636	.0470	8	2	1	11	.8000	.9167	.8563	.8544	.0119

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 15. Precision, recall, F_1-score, classification accuracy (%), AUC, confusion matrix of SVM (all the datasets)

Dataset Used	Class	Kernel	ϒ	Degree	C	Precision	Recall	F_1-Score	Accuracy (%)	AUC	Confusion Matrix
Breast Cancer	2	RBF	1	3	10	0.62	0.57	0.33	0.71	0.58	[[5 11] [2 12]]
		Polynomial	0.3	2	1	0.56	0.56	0.56	0.53	0.53	[[9 7] [7 7]]
Prostate Cancer	2	RBF	0.7	1	1	0.93	1	0.96	0.95	0.91	[[28 0] [2 10]]
		Polynomial	0.7	1	1	0.9	0.32	0.47	0.5	0.62	[[9 19] [1 11]]
Colon Tumor	2	RBF	1	2	1	0.68	1	0.81	0.7	0.57	[[13 0] [6 1]]
		Polynomial	0.3	2	1	0.62	0.76	0.68	0.55	0.45	[[10 3] [6 1]]
Leukemia	2	RBF	1	3	10	0.88	1	0.94	0.90	0.83	[[16 0] [2 4]]
		Polynomial	0.3	2	1	0.88	1	0.94	0.89	0.83	[[16 0] [2 4]]

Table 16. Testing Accuracy and AUC of all datasets

Data set	ELM		OSELM	WKELM		RKELM		SVM-RBF		SVM-POLY	
	Max. Tst Acc	AUC	Max. Tst Acc	Max. Tst Acc	AUC	Max. Tst Acc	AUC	Max. Tst Acc	AUC	Max. Tst Acc	AUC
Breast Cancer	.9000	.6923	.8088	.9667	.8902	.9333	.8253	0.71	0.58	0.53	0.53
Prostate Cancer	.8500	.6152	.7831	.9250	.8751	.9750	.7452	0.95	0.91	0.50	0.61
Colon Tumor	.9000	.6844	.7076	.9500	.7902	.9500	.7902	0.70	0.57	0.55	0.45
Leukemia	.8636	.6436	.6909	.9545	.8653	.8636	.8423	0.90	0.83	0.89	0.83

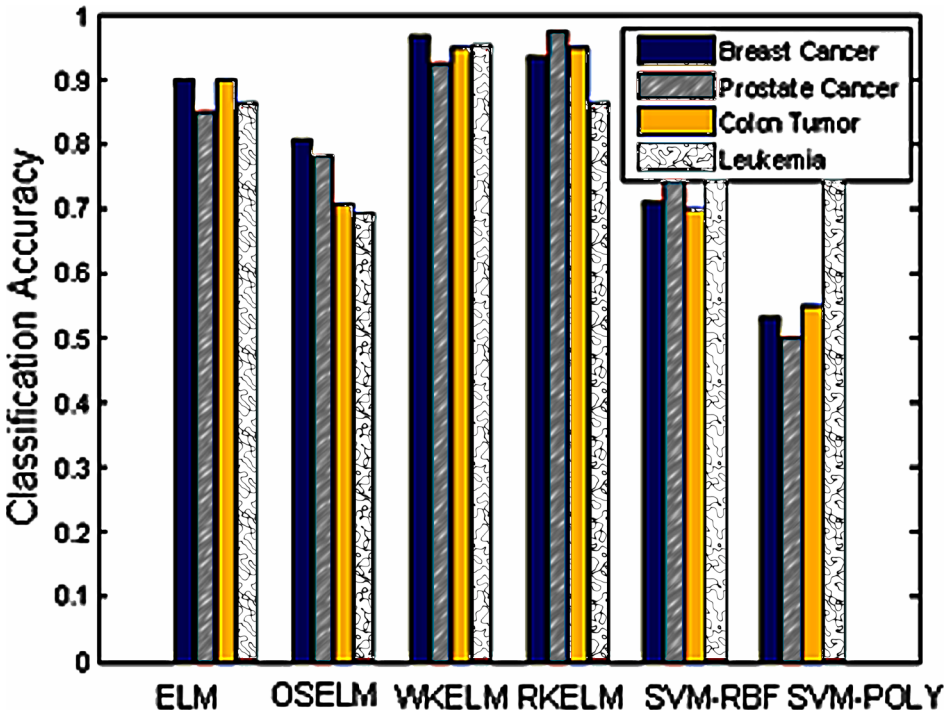
tained from all the datasets are at par with the contemporary researchers. For example in (Nguyen et al., 2015) the classification accuracy of breast cancer data using PNN, SVM, FARTMAP, ANFIS, SAM and GASAM models are 94, 96, 95, 96, 96 and 97 respectively.

8. CONCLUSION

Microarray data analysis and classification are very pertinent for the effective diagnosis of different varieties of cancer. However, microarray medical datasets exhibit high dimensionality thereby causing difficulties in accurately identifying the type of cancer. Thus the most relevant features are extracted from the data by using a meta-heuristic cat swarm optimization (CSO) technique. Further to improve the feature selection process a modified CSO known as MCSO is proposed in this paper and its performance

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 19. Testing accuracy for all models with all datasets



is compared with CSO validating its superiority over the latter. The selected features are used as inputs to ELM and its variants OSELM, WKELM and RKELM for microarray medical data classification. To provide comparison two powerful SVM classifiers known as SVMRBF and SVMPoly are used for classification of the microarray data. Various performance metrics like training and testing accuracy, sensitivity, specificity, confusion matrix, Gmean, F-score and the area under the receiver operating characteristic (ROC) curve, etc. are used to evaluate the efficacy of all models. From simulation experiments it can be concluded that the kernel ELM (KELM) outperforms ELM, OSELM and SVM computing models irrespective of the datasets used. However, when the comparison is made between two variants of KELM i.e. WKELM and RKELM, WKELM yields better accuracy in three out of four datasets used.

REFERENCES

Al-Ani, A., Alsukker, A., & Khushaba, R. N. (2013). Feature subset selection using differential evolution and a wheel based search strategy. *Swarm and Evolutionary Computation*, *9*, 15–26. doi:10.1016/j.swevo.2012.09.003

Banati & Bajaj. (2011). Firefly based feature selection approach. *International Journal of Computer Science Issues, 8*(4).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bencherif, M., Bazi, Y., Guessoum, A., Alajlan, N., Melgani, F., & AlHichri, H. (2015). Fusion of extreme learning machine and graph-based optimization methods for active classification of remote sensing images. *Geoscience and Remote Sensing Letters, IEEE*, *12*(3), 527–531. doi:10.1109/LGRS.2014.2349538

Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A., Benítez, J. M., & Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences*, *282*, 111–135. doi:10.1016/j.ins.2014.05.042

Bowes, D., Hall, T., & Gray, D. (2012). Comparing the performance of fault prediction models which report multiple performance measures: recomputing the confusion matrix. In *Proceedings of the 8th International Conference on Predictive Models in Software Engineering*. ACM. 10.1145/2365324.2365338

Chaudhuri, B. B., & Bhattacharya, U. (2000). Efficient training and improved performance of multilayer perceptron in pattern classification. *Neurocomputing*, *34*(1), 11–27. doi:10.1016/S0925-2312(00)00305-2

Chen, A. H., & Yang, C. (2012). The improvement of breast cancer prognosis accuracy from integrated gene expression and clinical data. *Expert Systems with Applications*, *39*(5), 4785–4795. doi:10.1016/j.eswa.2011.09.144

Chen, Z. (2013). Online sequential ELM based transfer learning for transportation mode recognition. *Cybernetics and Intelligent Systems (CIS), IEEE Conference on*. IEEE. 10.1109/ICCIS.2013.6751582

Chu, F., & Wang, L. (2005). Applications of support vector machines to cancer classification with microarray data. *International Journal of Neural Systems*, *15*(06), 475–484. doi:10.1142/S0129065705000396 PMID:16385636

Chu, S.-C., & Tsai, P.-W. (2007). Computational intelligence based on the behavior of cats. *International Journal of Innovative Computing, Information, & Control*, *3*(1), 163–173.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: A comparison of three data mining methods. *Artificial Intelligence in Medicine*, *34*(2), 113–127. doi:10.1016/j.artmed.2004.07.002 PMID:15894176

Diao, R., & Shen, Q. (2012). Feature selection with harmony search. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *42*(6), 1509–1523. doi:10.1109/TSMCB.2012.2193613 PMID:22645272

Díaz-Uriarte & De Andres. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics, 7*(1).

Duan, K.-B., Rajapakse, J. C., Wang, H., & Azuaje, F. (2005). Multiple SVM-RFE for gene selection in cancer classification with expression data. *NanoBioscience, IEEE Transactions on*, *4*(3), 228–234. doi:10.1109/TNB.2005.853657 PMID:16220686

Dunne, K., Cunningham, P., & Azuaje, F. (2002). Solutions to instability problems with sequential wrapper-based approaches to feature selection. *Journal of Machine Learning Research*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Elias, B. (2003). *Extracting landmarks with data mining methods*. COSIT. doi:10.1007/978-3-540-39923-0_25

Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J. W. W., Comber, H., & Bray, F. (2013). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *European Journal of Cancer*, *49*(6), 1374–1403. doi:10.1016/j.ejca.2012.12.027 PMID:23485231

Fernández-Navarro, F., Hervás-Martínez, C., Ruiz, R., & Riquelme, J. C. (2012). Evolutionary generalized radial basis function neural networks for improving prediction accuracy in gene classification using feature selection. *Applied Soft Computing*, *12*(6), 1787–1800. doi:10.1016/j.asoc.2012.01.008

Ferri, C., Hernández-Orallo, J., & Modroiu, R. (2009). An experimental comparison of performance measures for classification. *Pattern Recognition Letters*, *30*(1), 27–38. doi:10.1016/j.patrec.2008.08.010

Fong, S. (2013). Selecting Optimal Feature Set in High-Dimensional Data by Swarm Search. *Journal of Applied Mathematics*.

Furey, T. S., Cristianini, N., Duffy, N., Bednarski, D. W., Schummer, M., & Haussler, D. (2000). Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics (Oxford, England)*, *15*(410), 906–914. doi:10.1093/bioinformatics/16.10.906 PMID:11120680

Guo, G. (2003). *KNN model-based approach in classification. In On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE* (pp. 986–996). Springer Berlin Heidelberg. doi:10.1007/978-3-540-39964-3_62

Guyon, I., & Elisseeff, A. (2006). *An introduction to feature extraction. In Feature extraction* (pp. 1–25). Springer Berlin Heidelberg.

Hsu, Chang, & Lin. (2003). *A practical guide to support vector classification*. Academic Press.

Huang, C.-L., & Wang, C.-J. (2006). A GA-based feature selection and parameters optimizationfor support vector machines. *Expert Systems with Applications*, *31*(2), 231–240. doi:10.1016/j.eswa.2005.09.024

Huang, G.-B. (2012). Extreme learning machine for regression and multiclass classification. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *42*(2), 513–529. doi:10.1109/TSMCB.2011.2168604 PMID:21984515

Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2004). Extreme learning machine: a new learning scheme of feed forward neural networks. In *Proceedings, 2004 International Joint Conference on Neural Networks*. IEEE.

Huang, W. (2013). Liver tumor detection and segmentation using kernel-based extreme learning machine. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE.

Jakhar, Kaur, & Singh. (2011). Face recognition using bacteria foraging optimization-based selected features. *International Journal of Advanced Computer Science and Applications*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kanan, H. R., & Faez, K. (2008). An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system. *Applied Mathematics and Computation*, *205*(2), 716–725. doi:10.1016/j.amc.2008.05.115

Kanavos, P. (2006). The rising burden of cancer in the developing world. *Annals of Oncology, 17*(S8), 15-23.

Karg, M. (2009). A comparison of PCA, KPCA and LDA for feature extraction to recognize affect in gait kinematics. *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*. IEEE. 10.1109/ACII.2009.5349438

Khushaba, R. N., Al-Ani, A., & Al-Jumaily, A. (2008). Differential evolution based feature subset selection. *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE. 10.1109/ICPR.2008.4761255

Lee, Y., & Lee, C.-K. (2003). Classification of multiple cancer types by multicategory support vector machines using gene expression data. *Bioinformatics (Oxford, England)*, *19*(9), 1132–1139. doi:10.1093/bioinformatics/btg102 PMID:12801874

Li, B., Rong, X., & Li, Y. (2014). An Improved Kernel Based Extreme Learning Machine for Robot Execution Failures. *The Scientific World Journal*. PMID:24977234

Li, D.-C., Liu, C.-W., & Hu, S. C. (2011). A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets. *Artificial Intelligence in Medicine*, *52*(1), 45–52. doi:10.1016/j.artmed.2011.02.001 PMID:21493051

Li, D.-C., Liu, C.-W., & Hu, S. C. (2011). A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets. *Artificial Intelligence in Medicine*, *52*(1), 45–52. doi:10.1016/j.artmed.2011.02.001 PMID:21493051

Li, J., & Liu, H. (2002). *Kent ridge bio-medical data set repository*. Institute for Infocomm Research. Retrieved from http://sdmc.lit.org.sg/GEDatasets/Datasets.html

Lin, K.-C. (2014). Feature selection and parameter optimization of support vector machines based on modified cat swarm optimization. *International Journal of Distributed Sensor Networks*.

Lin, K.-C., & Chien, H.-Y. (2009). CSO-based feature selection and parameter optimization for support vector machine. *Pervasive Computing (JCPC), 2009 Joint Conferences on*. IEEE.

Lin, S.-W., Lee, Z.-J., Chen, S.-C., & Tseng, T.-Y. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied Soft Computing*, *8*(4), 1505–1512. doi:10.1016/j.asoc.2007.10.012

Liu, et al. (2004). A combinational feature selection and ensemble neural network method for classification of gene expression data. *BMC Bioinformatics, 5*(1).

Liu, Y., Wang, G., Chen, H., Dong, H., Zhu, X., & Wang, S. (2011). An improved particle swarm optimization for feature selection. *Journal of Bionics Engineering*, *8*(2), 191–200. doi:10.1016/S1672-6529(11)60020-6

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

López, M. M., Ramírez, J., Górriz, J. M., Álvarez, I., Salas-Gonzalez, D., Segovia, F., & Chaves, R. (2009). SVM-based CAD system for early detection of the Alzheimer's disease using kernel PCA and LDA. *Neuroscience Letters*, *464*(3), 233–238. doi:10.1016/j.neulet.2009.08.061 PMID:19716856

Malathi, V., Marimuthu, N., & Baskar, S. (2010). Intelligent approaches using support vector machine and extreme learning machine for transmission line protection. *Neurocomputing*, *73*(10), 2160–2167. doi:10.1016/j.neucom.2010.02.001

McCallum, A., & Nigam, K. (1998). A comparison of event models for naive bayes text classification. AAAI-98 workshop on learning for text categorization (vol. 752).

Misra, B. B., & Dehuri. (2007). *Functional Link Artificial Neural Network for Classification Task in Data Mining 1*. Academic Press.

Nakamura, R. et al. (2012). BBA: A binary bat algorithm for feature selection. *Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on*. IEEE.

Nguyen, T., Khosravi, A., Creighton, D., & Nahavandi, S. (2015). Classification of healthcare data using genetic fuzzy logic system and wavelets. *Expert Systems with Applications*, *42*(4), 2184–2197. doi:10.1016/j.eswa.2014.10.027

Panda, G., Mohan Pradhan, P., & Majhi, B. (2011). IIR system identification using cat swarm optimization. *Expert Systems with Applications*, *38*(10), 12671–12683. doi:10.1016/j.eswa.2011.04.054

Pedregosa, F. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *27*(8), 1226–1238. doi:10.1109/TPAMI.2005.159 PMID:16119262

Quinlan, J. (2014). *Ross. C4. 5: Programs for machine learning*. Elsevier.

Rodrigues, D. (2013). BCS: A binary cuckoo search algorithm for feature selection. *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE. 10.1109/ISCAS.2013.6571881

Saeys, Inza, & Larrañaga. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics, 23*(19), 2507-2517.

Saha, S. K., Ghoshal, S. P., Kar, R., & Mandal, D. (2013). Cat swarm optimization algorithm for optimal linear phase FIR filter design. *ISA Transactions*, *52*(6), 781–794. doi:10.1016/j.isatra.2013.07.009 PMID:23958491

Senliol, B. (2008). Fast Correlation Based Filter (FCBF) with a different search strategy. *Computer and Information Sciences, 2008. ISCIS'08. 23rd International Symposium on*. 10.1109/ISCIS.2008.4717949

Siegel, R. L., Miller, K. D., & Jemal, A. (2015). Cancer statistics, 2015. *CA: a Cancer Journal for Clinicians*, *65*(1), 5–29. doi:10.3322/caac.21254 PMID:25559415

Slaby, A. (2007). ROC analysis with MATLAB. In *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on*. IEEE. 10.1109/ITI.2007.4283768

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Statnikov, A., Tsamardinos, I., Dosbayev, Y., & Aliferis, C. F. (2005). GEMS: A system for automated cancer diagnosis and biomarker discovery from microarray gene expression data. *International Journal of Medical Informatics*, *74*(7), 491–503. doi:10.1016/j.ijmedinf.2005.05.002 PMID:15967710

Subashini, T. S., Ramalingam, V., & Palanivel, S. (2009). Vennila Ramalingam, and S. Palanivel. "Breast mass classification based on cytological patterns using RBFNN and SVM. *Expert Systems with Applications*, *36*(3), 5284–5290. doi:10.1016/j.eswa.2008.06.127

Subasi, A. (2004). Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Systems with Applications*, 1–11.

Subasi, A., & Ismail Gursoy, M. (2010). EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications*, *37*(12), 8659–8666. doi:10.1016/j.eswa.2010.06.065

Suresh, S., Saraswathi, S., & Sundararajan, N. (2010). S. Saraswathi, and N. Sundararajan. "Performance enhancement of extreme learning machine for multi-category sparse data classification problems. *Engineering Applications of Artificial Intelligence*, *23*(7), 1149–1157. doi:10.1016/j.engappai.2010.06.009

Tan, M. (2007). Comparative study of kernel based classification and feature selection methods with gene expression data. In *Computational methods of feature selection*. CRC Press.

Unler, A., & Murat, A. (2010). A discrete particle swarm optimization method for feature selection in binary classification problems. *European Journal of Operational Research*, *206*(3), 528–539. doi:10.1016/j.ejor.2010.02.032

Vert, J.-P., & Kanehisa, M. (2002). Graph-driven feature extraction from microarray data using diffusion kernels and kernel CCA. *Advances in Neural Information Processing Systems*.

Vlahou, A. (2003). Diagnosis of ovarian cancer using decision tree classification of mass spectral data. *BioMed Research International*, *2003*(5), 308–314. PMID:14688417

Zhang, et al. (2007). Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 4*(3), 485-495.

Zhang, Y., Ding, C., & Li, T. (2008). Gene selection algorithm by combining reliefF and mRMR. *BMC Genomics*, *9*(Suppl 2), S27. doi:10.1186/1471-2164-9-S2-S27 PMID:18831793

This research was previously published in the International Journal of Applied Evolutionary Computation (IJAEC), 7(3); edited by Wei-Chiang Samuelson Hong; ; pages 71-100; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 43
Integrated Approach for Automatic Crackle Detection Based on Fractal Dimension and Box Filtering

Cátia Pinho
University of Aveiro, Portugal

Cristina Jácome
University of Aveiro, Portugal

Ana Oliveira
University of Aveiro, Portugal

João Manuel Rodrigues
University of Aveiro, Portugal

Alda Marques
University of Aveiro, Portugal

ABSTRACT

Crackles are adventitious respiratory sounds (RS) that provide valuable information on different respiratory conditions. Nevertheless, crackles automatic detection in RS is challenging, mainly when collected in clinical settings. This study aimed to develop an algorithm for automatic crackle detection/characterisation and to evaluate its performance and accuracy against a multi-annotator gold standard. The algorithm is based on 4 main procedures: i) recognition of a potential crackle; ii) verification of its validity; iii) characterisation of crackles parameters; and iv) optimisation of the algorithm parameters. Twenty-four RS files acquired in clinical settings were selected from 10 patients with pneumonia and cystic fibrosis. The algorithm performance was assessed by comparing its results with a multi-annotator gold standard agreement. High level of overall performance (F-score=92%) was achieved. The results highlight the potential of the algorithm for automatic crackle detection and characterisation of RS acquired in clinical settings.

DOI: 10.4018/978-1-7998-1204-3.ch043

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Acute and chronic respiratory diseases are a major health, societal and economic burden worldwide (World Health Organization, 2008). Therefore, during the last decade, several research efforts have been dedicated to improve the diagnosis, management and monitoring of patients with respiratory diseases (Dinis, Campos, Rodrigues, & Marques, 2012).

Although several measures are available to diagnose and monitor respiratory diseases, conventional auscultation remains the most widely used. The stethoscope has been used for 200 years to perform clinical examinations as it is simple, non-invasive, economic, practical and useful in all populations and settings (Bohadana, Izbicki, & Kraman, 2014). However, it is known that conventional auscultation is subjective, i.e., the interpretation of respiratory sounds depends on the stethoscope properties (Welsby & Earis, 2001), hearing ability and clinical experience of users (Sovijärvi, Vanderschoot, & Earis, 2000) and their capacity to memorise sound patterns (e.g., number and characteristics of adventitious respiratory sounds) (Marques, Bruton, & Barney, 2006).

To overcome the subjectivity associated with conventional auscultation, research efforts have been devoted to improve computerised respiratory sound analysis (CORSA). CORSA consists of recording respiratory sounds with an electronic device and objectively analysing/classifying them based on advanced digital signal processing techniques (Sovijärvi, Vanderschoot, & Earis, 2000). The digitalization of respiratory sounds recordings has the advantages of: i) being compatible with digital signal processing techniques for sound analysis and classification; ii) allowing the precise definition of recording characteristics (e.g., recording level and bandwidth), which can be adjusted to the type of signals that are going to be studied (Cheetham, Charbonneau, Giordano, Helistö, & Vanderschoot, 2000); and iii) overcoming the barriers associated with clinicians having to memorise sound patterns. Through CORSA, respiratory sounds were found to be a more sensitive indicator, detecting and characterising the severity of respiratory diseases before any other measure (Gavriely, Nissan, Cugell, & Rubin, 1994). Furthermore, CORSA can also boost the development of mathematical models of the underlying physical mechanisms of respiratory sound production, to characterise the interaction of mechanical forces, airflow and sound transmission within the respiratory tract (Earis & Cheetham, 2000). This may provide important information to understand and relate different processes of diseases (Earis & Cheetham, 2000).

Computerised respiratory sounds are thus a simple, objective and non-invasive measure to assess the function of the respiratory system (Bohadana et al., 2014). Therefore, special attention has been given to their automatic detection and characterisation, as changes in their properties can early inform the presence of several respiratory conditions (Sovijärvi, A., Malmberg, L., et al., 2000), e.g., pneumonia (Piirila, 1992), bronchiectasis and cystic fibrosis (Marques, Bruton, & Barney, 2009).

Briefly, respiratory sounds can be classified into normal and adventitious respiratory sounds. Adventitious respiratory sounds are continuous (i.e., wheezes and rhonchus) or discontinuous (i.e., fine crackles and coarse crackles) sounds superimposed on normal respiratory sounds (Bohadana et al., 2014). The specific characteristics of wheezes, such as its duration (longer than 100 ms) and exhibition of distinct peaks in the frequency domain (> 100 Hz) have facilitated the development of methods for its automatic detection and characterisation (Taplidou & Hadjileontiadis, 2007). Crackles, however, are discontinuous, transient and explosive sounds of no more than 20 ms (Sovijärvi, A., Malmberg, L., et al., 2000). These characteristics introduced additional complexity in their automatic detection, and thus further research is needed to overcome this problem.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

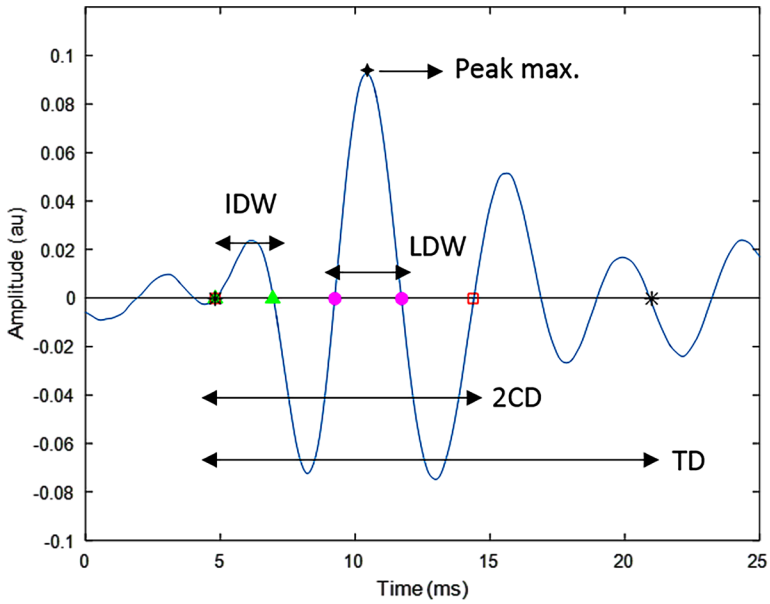
Crackles occur frequently in cardiorespiratory diseases (Charbonneau et al., 2000; Sovijärvi, A., Malmberg, L., et al., 2000) and their presence has been associated with the severity of diseases. Therefore, the study and detailed analysis of these discontinuous adventitious sounds is important to early detect respiratory diseases and enhance their assessment and monitoring. Previous research has predicted that crackles originating from smaller airways are shorter in duration (i.e., fine crackles) than those originating from larger airways (i.e., coarse crackles) (Sovijärvi, A., Malmberg, L., et al., 2000). Fine crackles are high-pitched inspiratory events that tend to occur in the mid-to-late inspiration and coarse crackles are low-pitched and often occur in the early inspiration and occasionally in expiration as well (Sovijärvi, A., Malmberg, L., et al., 2000). The most studied crackles parameters are:

- The initial deflection width (IDW), which is the duration of the first deflection in a crackle waveform;
- The largest deflection width (LDW), which is the duration of the deflection of the largest amplitude in a crackle waveform;
- The two cycle duration (2CD), which is the time from the beginning of the initial deflection of a crackle to the point where the waveform of the crackle has completed two cycles (Charbonneau et al., 2000; Sovijärvi, A., Malmberg, L., et al., 2000).

These and other parameters are represented schematically in Figure 1.

Despite the great value of crackles to contribute for the diagnosis and monitoring of respiratory diseases, their recognition and characterization in respiratory sound files recorded in clinical settings is particularly challenging. Concerning human detection, the level of agreement among health professionals' annotation of crackles has been reported has poor, independently of the health professional

Figure 1. Crackle time-domain parameters: Initial Deflection Width (IDW), Largest Deflection Width (LDW), Two-Cycle Duration (2CD), Total Duration (TD) and Maximum Peak (Peak max).



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

background and years of experience (Allingame, Williams, Jenkins, & Tucker, 1995; Brooks & Thomas, 1995; Machado, Oliveira, Aparício, & Marques, 2015). Computerised methods for automatic detection of crackles eliminate this subjectivity (Hadjileontiadis & Rekanos, 2003; Lu & Bahoura, 2008; Vannuccini, Rossi, & Pasquali, 1998). However, the performance of computerised methods is also uncertain, as respiratory sound files acquired in clinical settings have often artefacts (e.g., environmental noise and movement artefacts), which affect the recognition accuracy of algorithms (Chang & Lai, 2010).

Several signal processing techniques have been proposed, including digital filters (Ono, Arakawa, Mori, Sugimoto, & Harashima, 1989), spectrogram analysis (Kaisla et al., 1991), time-domain analysis (Kaisla et al., 1991), auto-regressive models (Hadjileontiadis & Panas, 1996) and wavelet-packet transform methods (Lu & Bahoura, 2008), fuzzy filters (Mastorocostas, Tolias, Theocharis, Hadjileontiadis, & Panas, 2000) and fractal dimension filtering (Hadjileontiadis & Rekanos, 2003). Despite the high values of sensitivity and specificity associated with these techniques, limited testing (Hadjileontiadis & Rekanos, 2003; Lu & Bahoura, 2008; Vannuccini, Rossi, & Pasquali, 1998) have been performed with respiratory sound files recorded in clinical settings and few have been validated against a multi-annotator gold standard (Quintas, Campos, & Marques, 2013). The gold standard annotation of respiratory sound files can only be achieved by human expert annotation (Marques et al., 2006). Nevertheless, human annotation has an underlying subjectivity (Marques et al., 2006) and therefore, manual annotation performed independently by multiple annotators is essential to produce a more accurate gold standard. Thus, the gold standard must result from the application of statistical agreement criteria to multiple independent annotations obtained for each file (Quintas et al., 2013). Therefore, the acquisition of respiratory sounds in clinical settings and the validation of automatic detection algorithms against a multi-annotator gold standard, are key features to address in the development of more robust algorithms.

Taking these factors into consideration, it may be hypothesised that, when tested in such conditions (i.e., sounds recorded in clinical settings and validation of results against a multi-annotator gold standard), the signal processing techniques proposed in the literature will have lower performances than originally reported. Therefore, a robust algorithm to detect and characterise crackles in respiratory sound files acquired in clinical settings is needed.

This study aimed to develop a new integrated approach for automatic crackle detection and characterisation (based on fractal dimension and box filtering techniques), and to evaluate its performance and accuracy against a multi-annotator gold standard (obtained by the agreement among three experts). Accordingly, this paper is organised in the following four sections: i) *Methods*, which describes the methodologies underlying the development and validation of the proposed algorithm; ii) *Results*, which presents the results of the algorithm's validation in terms of sensitivity (SE); positive predictive value (PPV) and F-score; iii) *Discussion*, which discusses the results and highlights possible improvements; and iv) *Conclusion and future work*, which concludes the paper and presents future work.

METHODS

The development and evaluation of the algorithm for automatic crackle detection and characterisation was performed using a sample of twenty-four respiratory sound files. Fifteen respiratory sound files were acquired in a community clinical setting from 6 adult patients with pneumonia (3 females; mean age 46 $\pm$ 14.6 years old) and 9 respiratory sound files were acquired in a hospital environment from 4 patients with cystic fibrosis (6 females; mean age 32.3 $\pm$ 18.1 years old). The selection of these respiratory sound

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

files was based in previous research showing that pneumonia and cystic fibrosis are respiratory diseases characterised by the presence of crackles (Marques et al., 2009; Piirila, 1992). These respiratory sound files were obtained from 2 academic repositories managed by one of the authors of this paper (Alda Marques). Respiratory sounds of patients with cystic fibrosis were collected during a PhD project at the School of Health Sciences of the University of Southampton (SFRH/BD/21375/2005). Respiratory sound files of patients with pneumonia belong to a repository built at the School of Health Sciences of the University of Aveiro in the scope of a research project (PTDC/SAU–BEB/101943/2008). Approvals for these studies were obtained from institutional ethic committees and respective national data protection.

All respiratory sounds were acquired following CORSA short-term sound acquisition guidelines (Rossi, Sovijärvi, Piirila, Vannuccini, & Dalmasso, 2000). After 5 min of quiet sitting, patients were instructed to breathe normally through the mouth while respiratory sounds were recorded.

Recordings were acquired three times in seven chest locations (trachea; left and right: anterior, lateral and posterior), for 25 seconds (Rossi et al., 2000). Two digital stethoscopes have been used: Thinklabs® digital stethoscope (Thinklabs® Rhythm: ds32a, Colorado, US) in patients with pneumonia, and with a WelchAllyn digital stethoscope (WelchAllyn Meditron, 5079-402) in patients with cystic fibrosis. Recordings were performed at a sampling rate (fs) of 44.1 kHz and saved in .wav format. Written informed consents were obtained before any data collection.

CRACKLES MANUAL MULTI-ANNOTATION

Firstly, each file duration was reduced to 10 seconds, because manual annotation is a time-consuming process, since there can be many crackles in a file of few seconds long (Dinis et al., 2012; Oliveira, A., Pinho, Dinis, Oliveira, & Marques, 2013). Then, all respiratory sound files were manually annotated by three respiratory researchers, with experience in visual-auditory crackle recognition, who independently identified the presence of crackles in each respiratory sound file. When the researcher considered that a crackle was present, he/she was requested to annotate the beginning and the end of the crackle. Respiratory Sound Annotation Software V1.1 was used to perform the manual annotation (Dinis et al., 2012).

For each respiratory sound file, a gold standard annotation was obtained by combining the annotations from the three researchers. An event was flagged as a crackle if at least two researchers had identified it, i.e., agreement by majority (Gross, Hadjileontiadis, Penzel, Koehler, & Vogelmeier, 2003). The agreement among the three researchers was set considering the maximum absolute peak within each crackle annotation (Quintas et al., 2013) and was assessed by dividing the number of crackles in which researchers agreed by the total number of crackles annotated. An agreement of 86% was achieved among the three researchers.

AUTOMATIC CRACKLE DETECTION

The proposed algorithm can be summarised in the following main steps:

1. Extraction of a window of interest of a potential crackle based on fractal dimension estimation (Hadjileontiadis & Rekanos, 2003; Katz, 1988; Lu & Bahoura, 2008) and box filtering techniques (Shen, Shen, Castan, & Zhang, 2002; Taplidou & Hadjileontiadis, 2007);

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. Verification of the validity of the potential crackle, considering CORSA established criteria (Charbonneau et al., 2000; Murphy, Del Bono, & Davidson, 1989);
3. Characterisation and extraction of crackle parameters;
4. Optimisation of algorithm's parameters.

To reduce computation requirements, the signal was down sampled to 11025 Hz (fs), after appropriate anti-aliasing filtering (i.e., a low-pass digital filter applied to significantly attenuate frequencies above 5512.5 Hz, which would otherwise interfere with the lower frequencies after down sampling). The signal filtering consisted in the application of a passband filter of [100 – 2000] Hz (finite impulse response, designed with a 83ms Blackman window), to eliminate high frequency noise, ensuring that the main features of the crackles were still preserved (Sovijärvi, A., Malmberg, L., et al., 2000; Vannuccini et al., 2000). A detailed description of each one of these steps is given below.

Extraction of a Window of Interest

The extraction of a window of interest identifying a possible crackle involved four different signal processing steps:

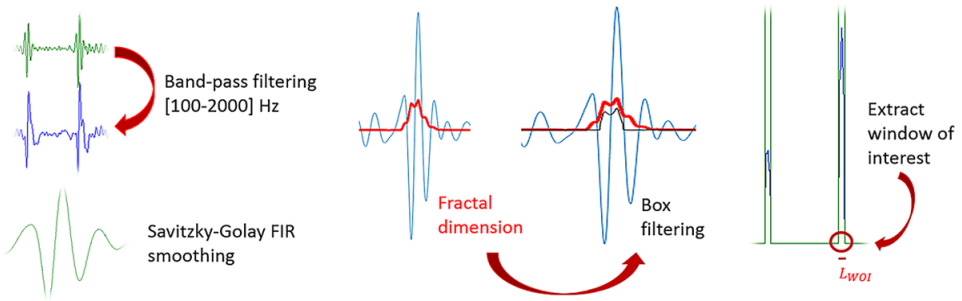
1. Savitzky-Golay (polynomial) finite impulse response (FIR) smoothing (Vannuccini et al., 1998);
2. Fractal dimension estimation (Hadjileontiadis & Rekanos, 2003; Katz, 1988; Lu & Bahoura, 2008);
3. Box filtering (Shen et al., 2002; Taplidou & Hadjileontiadis, 2007);
4. Application of a threshold to extract the beginning and the end of a window of interest.

A diagram summarising these steps is presented in Figure 2.

Firstly, a smoothing in the signal was performed to extract the remaining high frequency peaks of noise.

Secondly, the fractal dimension of the signal was estimated using a sliding window of length L_{FD}. Fractal dimension is a statistical measure that indicates an object's complexity in fractal geometry and it is used to evaluate the complexity of a waveform. The method proposed by Sevcik (2010) was used, as it is more sensitive to waveform changes and faster than other commonly used methods, such as the signals envelop method. Fractal dimension is defined by:

Figure 2. Main steps performed by the algorithm to extract a window of interest of length (L_{WOI}), i.e., identifying a potential crackle



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$FD = 1 + \frac{\ln(L)}{\ln(2N')} \tag{1}$$

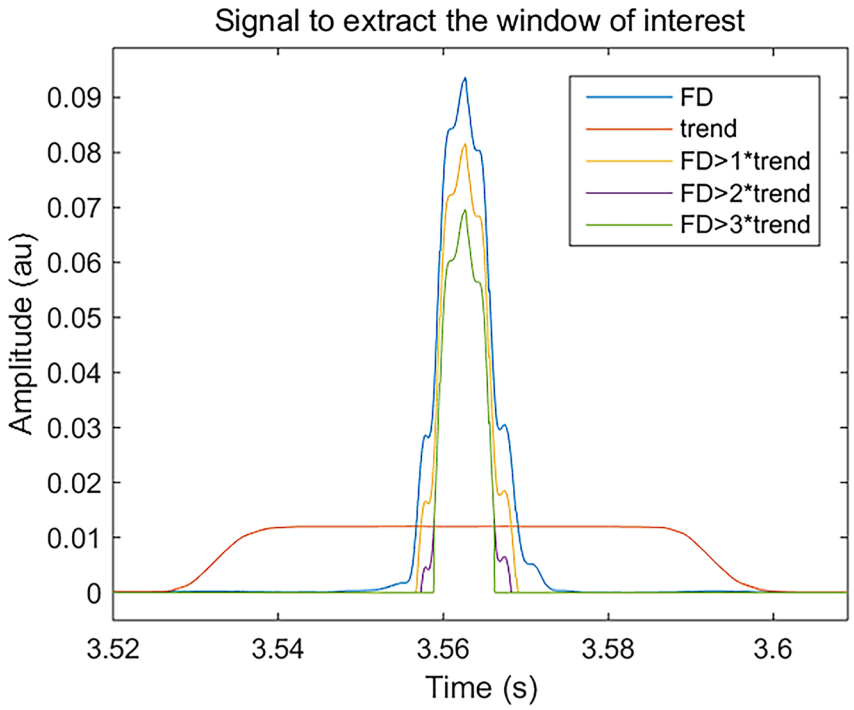
where N' is the number of steps in the waveform $\left(N' = L_{FD} \times fs - 1\right)$ and L represents the total length of the waveform, i.e., the sum of the Euclidean distance between successive data points:

$$L = \sum_{i=1}^{N'} dist(i, i+1) \tag{2}$$

Thirdly, an additional smoothing to estimate the trend of the fractal dimension signal in each time instance was applied. The estimation of the trend was based on box filtering, also known as average or mean filtering, with a sliding window of length L_{BF}. This is a commonly used technique to reduce noise, accomplished by replacing each amplitude sample with the average value of the surrounding samples, including itself. Therefore, the large amplitude variations between samples were removed and a much smoother signal was produced (i.e., the trend). This enhances the identification of crackles in noisy signals, such as sounds acquired in a clinical settings or sounds containing other adventitious respiratory sounds, e.g., wheezes.

In the fourth step, the fractal dimension signal was compared to an adaptive threshold to identify the window of interest of a potential crackle by *FD>X×trend*, where *X* is a free parameter for optimisation (Figure 3).

Figure 3. Adaptive threshold applied to the fractal dimension (FD) signal to identify the window of interest of a potential crackle by FD>X×trend, where X=1,2,3



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The final length of the window of interest was defined as the variable L_{WOI}.

Verification of Crackle

A function was developed to evaluate the validity of the potential crackle. This function intended to verify if the window of interest extracted met the CORSA established criteria for crackles detection (Kaisla et al., 1991). Also, additional considerations empirically established by the analysis of noisy signals were implemented to correct errors in the detection. Therefore, the following set of conditions were established:

1. The amplitude of the different peaks of the crackles had to be progressively lower than the LDW;
2. Peaks had to be progressively wider after the IDW (variable defined as W_{peaks});
3. Crackle zero-crossings were verified: minimum of 5 zero-crossings, to guarantee the calculation of 2CD, and maximum of 16 zero-crossings;
4. The mean absolute amplitude of the crackle had to be higher than F times the mean absolute amplitude of the background noise estimated from a segment of length L_{bckg} preceding the crackle (considering the CORSA criteria, that suggests that crackle amplitude have to be higher than 2*background noise);
5. Crackle IDW had to be higher than 1/8 of the LDW.

Characterisation and Extraction of Crackle Parameters

The following parameters were extracted to characterise the identified crackle: IDW, LDW and the 2CD (Charbonneau et al., 2000; Sovijärvi, A., Malmberg, L., et al., 2000). A detailed description of these crackle time-domain parameters is presented in the Figure 1 (please see the section *Introduction*). Furthermore, the classification of the crackle as fine or coarse was established, based on the criteria proposed by CORSA, i.e., a fine crackle was defined as having a 2CD $\leq$ 10 ms, and a coarse crackle has having a 2CD > 10 ms (Sovijärvi, A., Dalmasso, F., et al., 2000).

Optimisation of Algorithm's Parameters

Seven parameters were chosen to allow tuning the automatic crackle detection algorithm to improve performance. Three different values were tested for each parameter, resulting in a total of 2187 (3^7) combinations. The best combination was defined as the one achieving the highest F-score similarity with the gold standard annotation agreement.

The chosen tuning parameters and corresponding tested values were:

1. The width of the fractal window of approximately half the length of the crackle (Lu & Bahoura, 2008) (L_{FD} = *4, 6, 8 ms*);
2. The width of the box filtering window (L_{BF} = *40, 60, 80 ms*);
3. The *X* parameter applied in the adaptive thresholding step to identify the window of interest (*X = 1, 2, 3*);
4. The minimal length of the window of interest of each crackle (L_{WOI} = *0.5, 1, 2 ms*);
5. The widening tolerance range (= *1/5, 1/4, 1/3*) associated with the width of the peaks following IDW (W_{peaks} ± *widening tolerance* × W_{peaks});

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6. The length of the background sound before the crackle (L_{bckg} = *20, 40, all* ms*) *considering all background length before the crackle;
7. The factor multiplied by the amplitude of the background sound (*F = 1.6, 1.8, 2*).

Parameters 6 and 7 were established to evaluate the amplitude of the crackle when compared with the background noise (condition 'iv' defined in the section *Verification of crackle*). Only the background length before the crackle window was evaluated. This strategy was adopted to allow the extraction of crackles previously detected in that signal, ensuring that only the fluctuations of the background sound were considered and not the superposition of adventitious sounds (i.e., crackles).

Figure 4 provides the flow diagram of the proposed integrated approach to automatically detect crackles.

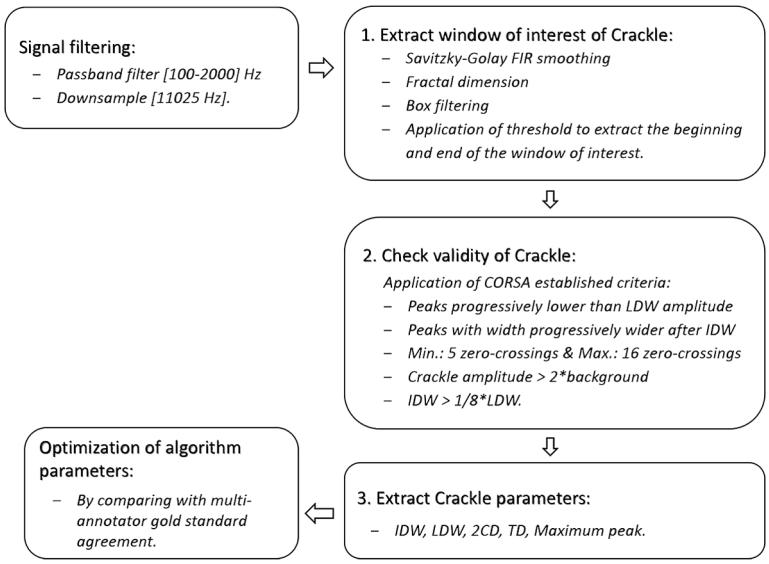
The algorithm regarding the proposed integrated approach was written in Matlab®R2009a (Mathworks, Natick, MA, USA).

CLASSIFICATION MEASURES TO VALIDATE THE ALGORITHM

The validation of the performance and accuracy of the automatic crackle detection algorithm was obtained by comparing the maximum absolute peak within each identified crackle with the results of the multi-annotator gold standard agreement, which for respiratory sounds is the judgment of respiratory experts (please see section *Crackles manual annotation*).

Formally, the proposed classification system is based in two classes (Cheetham et al., 2000), i.e., each event is mapped with a positive (P) (i.e., a crackle) or negative (N) (i.e., non-crackle) class label. The classification model then maps events to predicted classes (Fawcett, 2004, 2006). The event measure is given by the multi-annotator gold standard results, and the classifier refers to the algorithm outcomes. Thus, given a classifier and an event, it is possible to count four well-known parameters:

Figure 4. Flow diagram of the proposed algorithm to detect crackles



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. **True positive (TP):** If the event is flagged as crackle (i.e., positive) by the gold standard and classified as positive by the algorithm;
2. **False negative (FN):** If the event is flagged as positive by the gold standard and classified as negative (i.e., non-crackle) by the algorithm;
3. **True negative (TN):** If the event is flagged as negative by the gold standard and classified as negative by the algorithm;
4. **False positive (FP):** If the event is flagged as negative by the gold standard and classified as positive by the algorithm.

A two-by-two confusion matrix is used to describe this classification system and provide information regarding the actual and predicted classifications (Akay, 2009; Fawcett, 2006), Table 1.

This matrix forms the basis for many common metrics. Sensitivity (SE) and specificity (SP) are two measures that estimate a classifier's performance on different classes (Fawcett, 2006; Sokolova, Japkowicz, & Szpakowicz, 2006).

SE is defined as the ratio between crackles correctly detected (TP) and the sum of TP with the number of crackles not detected by the algorithm (FN), Equation (3):

$$SE = \frac{TP}{TP + FN} \tag{3}$$

SP is defined as the ratio between the events correctly classified as not a crackle (TN) and the sum of TN with the number of crackles incorrectly detected by the algorithm (FP), Equation (4):

$$SP = \frac{TN}{TN + FP} \tag{4}$$

As the analysed files are composed by 110250 points (i.e., $10s \times 11025s^{-1}$), even for a maximum number of crackles (e.g., 200), the order of magnitude of TN will be around 500 times above the FP count, and thus the SP value will always be approximately 100%. Thus, SP metrics were not considered in the results.

Other classification metrics is the positive predictive value (PPV) (Akay, 2009; Fawcett, 2006). PPV is the ratio between TP and the total number of crackles detected by the algorithm, i.e., correctly detected (TP) and incorrectly detected (FP), Equation (5):

Table 1. Confusion matrix representation

	Predicted	
Actual	**Positive**	**Negative**
Positive	True positive (**TP**)	False negative (**FN**)
Negative	False positive (**FP**)	True negative (**TN**)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$PPV = \frac{TP}{TP + FP} \tag{5}$$

The total performance F-score (F) is a multi-class classification metrics that combines PPV and SE, presenting, therefore, a more robust measure of the algorithm's performance, see Equation (6):

$$F = \frac{\left(1+\beta^2\right) \times SE \times PPV}{\left(\beta^2 \times PPV\right) + SE} \tag{6}$$

The parameter β allows to weight more heavily SE or PPV, i.e., $\beta > 1$ favours PPV and $\beta < 1$ SE (Hripcsak & Rothschild, 2005; Sokolova et al., 2006). In the F-score defined in the proposed classification system was assumed $\beta = 1$, i.e., an evenly balance between SE and PPV (Hripcsak & Rothschild, 2005; Sokolova et al., 2006), Equation (7):

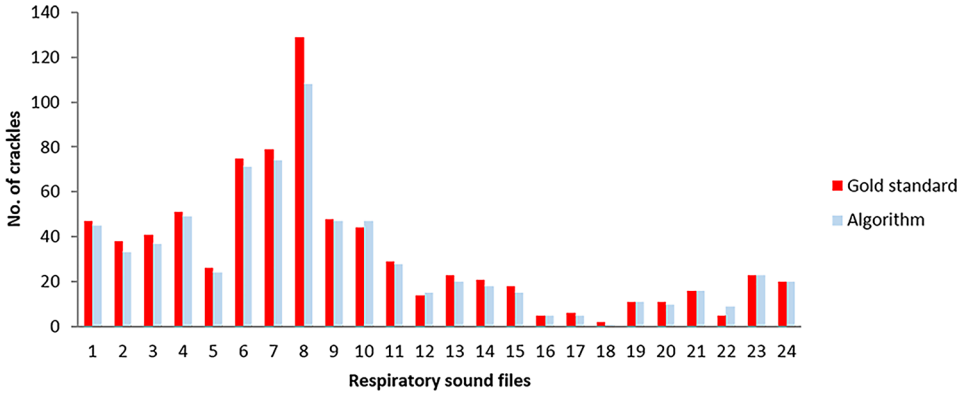
$$F = 2 \times \frac{SE \times PPV}{SE + PPV} \tag{7}$$

RESULTS

From a total of 2187 combinations of the threshold parameters, the performance of the automatic crackle detection algorithm was optimised with the following set of parameters: $L_{FD} = 6\ ms$; $L_{BF} = 60\ ms$; $X = 3$; $L_{WOI} = 2\ ms$; *widening tolerance = 1/4*; $L_{bckg} = 20\ ms$ and $F = 2$. A summary of the results obtained with this set of parameters is presented in Table 2.

The number of crackles identified with the optimised algorithm (from 1 to 108) was similar to the one obtained by the multi-annotator gold standard (from 2 to 129). These results, per respiratory sound file, are shown in Figure 5.

Figure 5. Crackles detected by manual annotation (multi-annotator gold standard agreement) and by the proposed algorithm for each respiratory sound file



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Results obtained for the 24 respiratory sound files analysed

N. File	Algorithm	GS	TP	FP	FN	SE	PPV	F
1	45	47	44	1	3	0.94	0.98	0.96
2	33	38	33	0	5	0.87	1.00	0.93
3	37	41	36	1	5	0.88	0.97	0.92
4	49	51	48	1	3	0.94	0.98	0.96
5	24	26	24	0	2	0.92	1.00	0.96
6	71	75	69	2	6	0.92	0.97	0.95
7	74	79	73	1	6	0.92	0.99	0.95
8	108	129	108	0	21	0.84	1.00	0.91
9	47	48	46	1	2	0.96	0.98	0.97
10	47	44	43	4	1	0.98	0.91	0.95
11	28	29	27	1	2	0.93	0.96	0.95
12	15	14	14	1	0	1.00	0.93	0.97
13	20	23	20	0	3	0.87	1.00	0.93
14	18	21	18	0	3	0.86	1.00	0.92
15	15	18	15	0	3	0.83	1.00	0.91
16	5	5	5	0	0	1.00	1.00	1.00
17	5	6	5	0	1	0.83	1.00	0.91
18	1	2	1	0	1	0.50	1.00	0.67
19	11	11	11	0	0	1.00	1.00	1.00
20	10	11	9	1	2	0.82	0.90	0.86
21	16	16	15	1	1	0.94	0.94	0.94
22	9	5	4	5	1	0.80	0.44	0.57
23	23	23	23	0	0	1.00	1.00	1.00
24	20	20	18	2	2	0.90	0.90	0.90

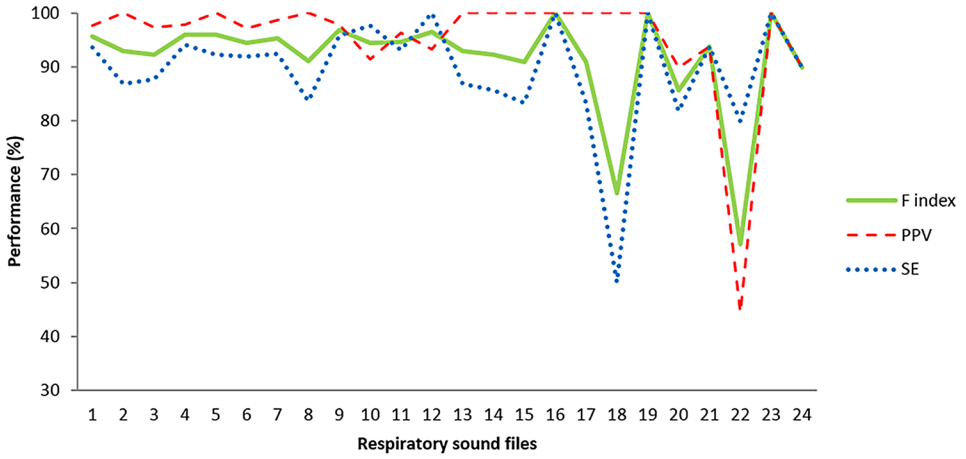
Note: N. File: number of sound files analysed; Algorithm: number of crackles detected by the proposed automatic crackle detection algorithm; GS – gold standard: counts of crackles identified by the multi-annotator gold standard agreement; TP – true positive: counts of crackles correctly detected; FP – false positive: counts of crackles incorrectly detected; FN – false negative: counts of non-crackles incorrectly identified; SE – sensitivity, PPV – positive predictive value, and F – F-score, present the results obtained for the three different classification metrics assessed to validate the algorithm performance.

Figure 6 presents the SE, PPV and F-score of the algorithm for each respiratory sound file. It can be observed that files 18 and 22, acquired in patients with pneumonia, were those in which the algorithm presented the lowest performance, F-score of 57% and 67% respectively.

However, when considering the 24 respiratory sound files, the average of SE, PPV and F-score of the proposed algorithm were considerably high (SE=89 ± 10%, PPV=95 ± 11% and F-score=92 ± 10%).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Sensitivity (SE), Positive Predictive Value (PPV) and F-score of the proposed algorithm for each respiratory sound



DISCUSSION

Crackles are adventitious respiratory sounds that frequently occur in cardiorespiratory diseases (Charbonneau et al., 2000; Sovijärvi, A., Malmberg, L., et al., 2000). Thus, their automatic detection and detailed analysis is of major importance for the diagnosis and monitoring of respiratory diseases. Nevertheless, the techniques provided in the literature for this calculus, e.g., Hadjileontiadis & Rekanos, 2003; Lu & Bahoura, 2008; Vannuccini et al., 1998, are poorly tested in clinical settings and not validated against a multi-annotator gold standard agreement.

Therefore, an integrated approach for automatic crackle detection has been proposed in this study to overcome the identified problems, based on:

1. The use of the combination of fractal dimension estimation and box filtering techniques to enhance the recognition of crackles in noisy signals;
2. The establishment of a set of rules to verify the validity of the detected crackle, grounded by CORSA established criteria (Charbonneau et al., 2000; Murphy et al., 1989);
3. The use for a multi-annotator gold standard manual annotation to validate the performance of the proposed automatic detection algorithm;
4. The optimisation of seven parameters established to improve the performance of the automatic crackle detection algorithm.

The results obtained from the optimisation of the seven parameters tested in a sample of 24 respiratory sound files acquired in clinical settings, demonstrated that this was an efficient and robust method for crackle detection/classification and highlights its potential to be used in respiratory sound files acquired in clinical practice.

The conditions initially implemented in the algorithm strictly followed the ones proposed by the CORSA criteria. However, during the iterative assessment of the algorithm, a sharp divergence between the results obtained with the algorithm and those from the gold standard was verified. To address this

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

problem, and as CORSA criteria are not updated since 2000, some refinements of the standard rules were implemented. For example, in the criteria stating that the width of the peaks should be progressively wider after IDW, a deviation of 25% was allowed. It should be noted, that despite the high subjectivity associated with human detection of crackles, this is still considered the only valid method for its detection (Marques et al., 2006), as health professionals are the ones who use it routinely to establish diagnosis and monitoring patients. Therefore, from the authors' point of view, algorithms should be developed to always match human gold standard annotation and not the opposite. Though, as human annotation has an underlying subjectivity the gold standard annotation should be provided by the statistical agreement criteria of multiple independent annotations (i.e., three or more) obtained for each sound file. Thus, in this study a multi-annotator gold standard agreement from three independently manual annotation experts was established.

The SE (89%), PPV (95%) and F-score (92%) of the proposed algorithm are comparable or even higher than those of other previously published methods (SE 80–91%; PPV 83−88%; F-score 86.7%) (Quintas et al., 2013; Vannuccini et al., 1998). However, it should be noted that, contrarily to other algorithms, the presented one was tested with respiratory sound files recorded in clinical settings, which further increase its potential to be used by health professionals in their daily practice. Only in two respiratory sound files from patients with pneumonia, the performance of the algorithm did not reach these high standard values (F-score of 57% and 67%). Such findings are related with the low number of crackles presented (TP) in these two respiratory sound files (1−9 crackles in the sound files), causing the few FP and FN to have a negative impact in the F-score. Nevertheless, this finding may not be clinically relevant as it has been reported that even healthy people present approximately 4 crackles per breathing cycle (Oliveira, Ana & Marques, 2014). Hence, errors of this magnitude might not be clinically significant for diagnosis and monitoring of respiratory diseases.

LIMITATIONS

This study has some limitations that need to be acknowledged. The proposed algorithm was developed by comparing the results against a multi-annotator gold standard obtained from the annotations of three respiratory researchers. Despite the high agreement achieved (86%), it is well-known that human annotation is associated with high levels of subjectivity, thus future studies should consider creating a repository of respiratory sounds annotated by an additional number of experts (e.g., five or more) to minimise bias. Also, a small sample of respiratory sounds files was included from adult patients with pneumonia and cystic fibrosis. Therefore, it would be a great asset to validate the proposed algorithm in large sets of data from patients with other respiratory diseases and different age ranges (young children and infants). Finally, although the proposed algorithm is currently processing in real-time, making it suitable to be used in the clinical practice, it would benefit from being included in Clinical Decision Support Systems. This would allow comparisons between patterns of samples and provide info/pre-diagnosis to the caregiver (Musen, Middleton, & Greenes, 2014), which is not addressed in the current version of the algorithm.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION AND FUTURE WORK

The proposed integrated approach for automatic crackle detection achieved a high performance (F-score 92%). This promising result highlights the potential of this new approach for automatic crackle detection and classification in respiratory sounds acquired in clinical settings. Future research could also test the developed algorithm in home-based environments and to be used in patients' tele-monitoring devices.

It would also be of great interest to develop a portable technology including a sound acquisition device plus an open source software integrating the developed algorithm. This would allow health professionals to have instant feedback on their respiratory sound recordings. Additionally, international collaborations are being established to acquire a large database of normal and adventitious respiratory sounds (including crackles, wheezes, rhonchi, stridor, etc.) that could be used to find unsuspected relationships between normal/adventitious respiratory sounds and respiratory diseases, summarise respiratory data into patterns and thus be used in Clinical Decision Support Systems.

ACKNOWLEDGMENT

This work was funded by Fundação para a Ciência e a Tecnologia, Portugal -Project PTDC/SAU-BEB/101943/2008, by the ERDF -European Regional Development Fund through the COMPETE Program (operational program for competitiveness) (ref. FCOMP-01-0124-FEDER-010919) and the National Strategic Reference Framework (ref. BI/UI64/3902/2010). The authors would also like to acknowledge João Dinis for his inputs in the development of the algorithm.

REFERENCES

Akay, M. F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Systems with Applications*, *36*(2 Part 2), 3240–3247. doi:10.1016/j.eswa.2008.01.009

Allingame, S., Williams, T., Jenkins, S., & Tucker, B. (1995). Accuracy and reliability of physiotherapists in the interpretation of tape-recorded lung sounds. *The Australian Journal of Physiotherapy*, *41*(3), 179–184. doi:10.1016/S0004-9514(14)60428-1 PMID:25026041

Bohadana, A., Izbicki, G., & Kraman, S. S. (2014). Fundamentals of lung auscultation. *The New England Journal of Medicine*, *370*(8), 744–751. doi:10.1056/NEJMra1302901 PMID:24552321

Brooks, D., & Thomas, J. (1995). Interrater reliability of auscultation of breath sounds among physical therapists. *Physical Therapy*, *75*(12), 1082–1088. PMID:7501711

Chang, G. C., & Lai, Y. F. (2010). Performance evaluation and enhancement of lung sound recognition system in two real noisy environments. *Computer Methods and Programs in Biomedicine*, *97*(2), 141–150. doi:10.1016/j.cmpb.2009.06.002 PMID:19615782

Charbonneau, G., Ademovic, E., Cheetham, B., Malmberg, L., Vanderschoot, J., & Sovijärvi, A. (2000). Basic techniques for respiratory sound analysis. *European Respiratory Review*, *10*(77), 625–635.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cheetham, B., Charbonneau, G., Giordano, A., Helistö, P., & Vanderschoot, J. (2000). Digitization of data for respiratory sound recordings. *European Respiratory Review*, *10*(77), 621–624.

Dinis, J., Campos, G., Rodrigues, J., & Marques, A. (2012). Respiratory Sound Annotation Software. *Paper presented at the International Conference on Health Informatics*, Vilamoura, Portugal.

Earis, J., & Cheetham, B. (2000). Future perspectives for respiratory sound research. *European Respiratory Review*, *10*, 641–646.

Fawcett, T. (2004). ROC Graphs: Notes and Practical Considerations for Researchers. Machine Learning, 31(1), 1-38.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. doi:10.1016/j.patrec.2005.10.010

Gavriely, N., Nissan, M., Cugell, D. W., & Rubin, A. H. (1994). Respiratory health screening using pulmonary function tests and lung sound analysis. *The European Respiratory Journal*, *7*(1), 35–42. doi:10.1183/09031936.94.07010035 PMID:8143829

Gross, V., Hadjileontiadis, L. J., Penzel, T., Koehler, U., & Vogelmeier, C. (2003). Multimedia Database "Marburg Respiratory Sounds (MARS)". *Paper presented at the 25ª Annual International Conference of the IEEE EMBS*, Cancun, Mexico.

Hadjileontiadis, L. J., & Panas, S. M. (1996, October 31 – November 3). Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics. Paper presented at the Engineering in Medicine and Biology Society.

Hadjileontiadis, L. J., & Rekanos, I. T. (2003). Detection of explosive lung and bowel sounds by means of fractal dimension. *Signal Processing Letters, IEEE*, *10*(10), 311–314. doi:10.1109/LSP.2003.817171

Hripcsak, G., & Rothschild, A. S. (2005). Agreement, the F-Measure, and Reliability in Information Retrieval. *Journal of the American Medical Informatics Association: JAMIA*, *12*(3), 296–298. doi:10.1197/jamia.M1733 PMID:15684123

Kaisla, T., Sovijärvi, A., Piirila, P., Rajala, H. M., Haltsonen, S., & Rosqvist, T. (1991). Validated Method for Automatic Detection of Lung Sound Crackles. *Medical & Biological Engineering & Computing*, *29*(5), 517–521. doi:10.1007/BF02442324 PMID:1817215

Katz, M. J. (1988). Fractals and the analysis of waveforms. *Computers in Biology and Medicine*, *18*(3), 145–156. doi:10.1016/0010-4825(88)90041-8 PMID:3396335

Lu, X. G., & Bahoura, M. (2008). An integrated automated system for crackles extraction and classification. *Biomedical Signal Processing and Control*, *3*(3), 244–254. doi:10.1016/j.bspc.2008.04.003

Machado, A., Oliveira, A., Aparício, J., & Marques, A. (2015). Respiratory auscultation: (dis)agreement between health professionals. *European Respiratory Journal, 46*(Suppl. 59).

Marques, A., Bruton, A., & Barney, A. (2006). Clinically useful outcome measures for physiotherapy airway clearance techniques: A review. *The Physical Therapy Review*, *11*(4), 299–307. doi:10.1179/108331906X163441

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Marques, A., Bruton, A., & Barney, A. (2009). Reliability of lung crackle characteristics in cystic fibrosis and bronchiectasis patients in a clinical setting. *Physiological Measurement*, *30*(9), 903–912. doi:10.1088/0967-3334/30/9/003 PMID:19641235

Mastorocostas, P. A., Tolias, Y. A., Theocharis, J. B., Hadjileontiadis, L. J., & Panas, S. M. (2000). An orthogonal least squares-based fuzzy filter for real-time analysis of lung sounds. *IEEE Transactions on Bio-Medical Engineering*, *47*(9), 1165–1176. doi:10.1109/10.867921 PMID:11008417

Murphy, R. L. Jr, Del Bono, E. A., & Davidson, F. (1989). Validation of an automatic crackle (rale) counter. *The American Review of Respiratory Disease*, *140*(4), 1017–1020. doi:10.1164/ajrccm/140.4.1017 PMID:2802365

Musen, M., Middleton, B., & Greenes, R. (2014). Clinical Decision-Support Systems. *Biomedical Informatics*, *46*(5), 643–674. doi:10.1007/978-1-4471-4474-8_22 PMID:23707417

Oliveira, A., & Marques, A. (2014). Respiratory sounds in healthy people: A systematic review. *Respiratory Medicine*, *108*(4), 550–570. doi:10.1016/j.rmed.2014.01.004 PMID:24491278

Oliveira, A., Pinho, C., Dinis, J., Oliveira, D., & Marques, A. (2013). Automatic Wheeze Detection and Lung Function Evaluation: a preliminary study. *Paper presented at the Healthinf - 6th International Conference on Health Informatics*, Barcelona.

Ono, M., Arakawa, K., Mori, M., Sugimoto, T., & Harashima, H. (1989). Separation of Fine Crackles from Vesicular Sounds by a Nonlinear Digital-Filter. *IEEE Transactions on Bio-Medical Engineering*, *36*(2), 286–291. doi:10.1109/10.16477 PMID:2645207

Piirila, P. (1992). Changes in crackle characteristics during the clinical course of pneumonia. *Chest*, *102*(1), 176–183. doi:10.1378/chest.102.1.176 PMID:1623749

Quintas, J., Campos, G., & Marques, A. (2013). Multi-algorithm respiratory crackle detection. *Paper presented at the Healthinf - 6th International Conference on Health Informatics*, Barcelona.

Rossi, M., Sovijärvi, A., Piirila, P., Vannuccini, L., & Dalmasso, F. V. J. (2000). Environmental and subject conditions and breathing manoeuvres for respiratory sound recordings. *European Respiratory Review*, *10*(77), 611–615.

Sevcik, C. (2010). A procedure to Estimate the Fractal Dimension of Waveforms. *Complexity International*, 5.

Shen, J., Shen, W., Castan, S., & Zhang, T. X. (2002). Sum-box technique for fast linear filtering. *Signal Processing*, *82*(8), 1109–1126. doi:10.1016/S0165-1684(02)00243-8

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In *Advances in Artificial Intelligence, LNCS* (Vol. 4304, pp. 1015-1021).

Sovijärvi, A., Dalmasso, F., Vanderschoot, J., Malmberg, P., Righini, G., & Stoneman, T. (2000). Definition of terms for applications of respiratory sounds. *European Respiratory Review*, *10*(77), 597–610.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sovijärvi, A., Malmberg, L., Charbonneau, G., Vanderschoot, J., Dalmasso, F., Sacco, C., ... Earis, J. (2000). Characteristics of breath sounds and adventitious respiratory sounds. *European Respiratory Review*, *10*(77), 591–596.

Sovijarvi, A., Vanderschoot, J., & Earis, J. (2000). Computerized respiratory sound analysis (CORSA).

Sovijärvi, A. R. A., Vanderschoot, J., & Earis, J. E. (2000). Standardization of computerized respiratory sound analysis. *European Respiratory Review*, *10*(77), 585.

Taplidou, S. A., & Hadjileontiadis, L. J. (2007). Wheeze detection based on time-frequency analysis of breath sounds. *Computers in Biology and Medicine*, *37*(8), 1073–1083. doi:10.1016/j.compbiomed.2006.09.007 PMID:17113064

Vannuccini, L., Earis, J., Helistö, P., Cheetham, B., Rossi, M., Sovijärvi, A., & Vanderschoot, J. (2000). Capturing and preprocessing of respiratory sounds. *European Respiratory Review*, *10*(77), 616–620.

Vannuccini, L., Rossi, M., & Pasquali, G. (1998). A new method to detect crackles in respiratory sounds. *Technology and Health Care*, *6*(1), 75–79. PMID:9754686

Welsby, P., & Earis, J. (2001). Some high pitched thoughts on chest examination. *Postgraduate Medical Journal*, *77*(912), 617–620. doi:10.1136/pmj.77.912.617 PMID:11571368

World Health Organization. (2008). *The World Health Report 2008: Primary Health Care - Now more than ever*. Geneva: Switzerland World Health Organization.

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 5(4); edited by Anastasius Moumtzoglou; pages 34-50; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 44
Electromyography-Based Functional Electrical Stimulation (FES) in Rehabilitation

Poulami Ghosh
Jadavpur University, India

Ankita Mazumder
Jadavpur University, India

Anwesha Banerjee
Jadavpur University, India

D.N. Tibarewala
Jadavpur University, India

ABSTRACT

Loss or impairment in the ability of muscle movement or sensation is called Paralysis which is caused by disruption of communication of nerve impulses along the pathway from the brain to the muscles. One of the principal reasons causing paralysis is Spinal Cord Injury (SCI) and Neurological rehabilitation by using neuro-prostheses, based on Functional Electrical Stimulation (FES) is extensively used for its treatment. Impaired muscles are activated by applying small amplitude electrical current. Electromyography (EMG), the recording of biosignals generated by muscle activity during the application of FES can be used as the control signal for FES based rehabilitative devices. This method is predominantly used for restoring upper extremity functioning (wrist, hand, elbow, etc.), standing, walking (speed, pattern) in stroke patients. FES, collaborated with conventional methods, has the potential to be utilized as a useful tool for rehabilitation and restoration of muscle strength, metabolic responses etc. in paralyzed patients.

DOI: 10.4018/978-1-7998-1204-3.ch044

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The EMG signal is produced by active contraction of the muscles. So, since electromyogram shows some function of muscle states it has been long harnessed for being used as a feedback signal for functional electrical stimulation or FES. Although there had been innumerable works on EMG for normal muscle conditions, less is known about electrically stimulated muscle EMG. (N Chesler et al, 1997)

Functional electrical stimulation can be controlled by signals obtained from surface EMG of activated voluntary muscles. EMG does not only signals muscular FES but also neuromuscular FES (Baldi, J. C. et al., 1998). It is so done that the recording electrodes are placed properly in position with respect to the stimulation electrodes. After that required processing is done on those recorded signals so as to minimize the artifacts and non-voluntary M-wave contribution. (B Daruish et al, 1998) It poses added difficulties on detecting EMG signals in the presence of FES as the FES electrode signals are of much higher magnitude as compared to the muscle generated signals and hence can result in saturating the biopotential amplifiers for a significant amount of time (Durfee, 1995). Hence many researchers earlier had considered removal of stimulation artifacts. But with newer researches this theory is abandoned and instead control scheme for FES is developed. Thus many studies have been carried out on optimal compromisation between voluntary muscle activity detection and artifact removal. (C Robinson et al, 1997)

The desired characteristics of the EMG electrodes are low noise, high impedance, high common mode rejection ratio preamplifier having typically a gain around 100 and a frequency range of 10 to 10,000 Hz. The preamplifier signal is then usually fed to a high pass filter, full wave rectifier and a blanking device that is used to reject the aforesaid stimulation artifacts. The signal now obtained is thus used a trigger signal for initiating or stopping stimulation (A Demosthenous et al, 2008)

Artificial neural network can be also an good option for eliminating adaptive interference effectively. Since the last two decades there had been studies over EMG controlled FES using ANN (artificial neural network) for pattern recognition and discrimination of EMG signals which in turn is used to control the FES. (Z Li et al, 2012) Prior to nineties there had been various studies on ANN implementation on EMG but all those were related to artificial limbs but not functional electrical stimulation. Controlling FES with the same idea was a difficult challenge as FES involved the controlling of peripheral nerves. Back propagation theory of Artificial neural network was used initially to answer this problem which was then modified with time. (H Yamaguchi et al, 2000)

In FES applications EMG are particularly used in those situations where an indication for muscle fatigue is needed. (Durfee 1995) Hence it can massively help in detecting muscle collapse in cases where FES is applied to treat hemi or paraplegic patients(Graupe et al, 1982) or those having ambulation aids or patient suffering from stroke(I Jerman et l, 2009) or spinal cord injury, (Hamzaid et al, 2009) there are quite a few methods to track fatigues but the most common fatigue tracking performance indicators are derived from the root mean square of electromyogram signal amplitude and from the MF i.e. the median frequency of EMG power spectral content. It has been found that such results indicated that for practical FES application it is pretty difficult to obtain proper fatigue tracking indicators.(J winslow et al, 2003)Thus it is very evident that EMG controlled FES has a very promising field of application and has multiple usages for varied kind of human diseases.

Apart from EMG signals, plenty of other signals get generated due the different activities going on continuously in the human body, known as Biosignals (Kaniusas, E. et al., 2012). These signals are basically used as the control signals needed to drive the FES devices. In case of feedback systems, these feedback signals are also extracted from these biosignals. The biosignals available with the human body

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

system are mainly Electrocardiography (signals generated from heart) (Fye, W. B., 1994), Electroencephalography (signals generated from brain) (Niedermeyer, E.et al., 2005), Electromyography (signals generated from muscles) (Oh, S. J. 2003), Electrooculography (signals generated from eye movement) (Miyake, Y. 2006), Electroneurography (signals generated from nerves) (Smorto, M. P. et al., 1974) and likewise. Few of these signals have been employed to drive the FES devices besides EMG signals like ENG, EEG. But EMG based FES is the most common among all the available methods. Other methods of controlling FES are discussed in details later in this chapter.

FUNCTIONAL ELECTRICAL STIMULATION (FES)

Electricity plays a major role in driving the nervous system of the human body. Whenever a cell or tissue gets stimulated, it becomes electrically activated and generates a very short electrical impulse known as action potential (Kandel, E et al., 2000). This action potential then stimulates the cell adjacent to it and then the subsequent cells and as a result the input stimulus gets propagated through the nervous system (Johnston, D., 1995). Apart from that, the brain also sends electrical pulses to muscles through nervous system for making the relevant muscles contract (Ebashi, S et al., 1969). These contractions are basically responsible for movement of different body parts. The idea of using electrical pulses for clinical application was first proposed by J.A. McWilliam in 1899 (first artificial pacemaker). After that a number of research works were executed for applying this idea for different types of clinical uses. W.T. Liberson et al. (1961) devised a stimulator for the prevention of drop foot in hemiplegic patients. This work evoked the idea of a new method of rehabilitation within scientists, known as Functional Electrical Stimulation (Rushton, D. N. 1997). This is a technique of stimulating inactive muscles by applying electrical pulses to it. FES is mainly applied to those muscles which are uninjured themselves and excitable but does not have an uninterrupted neural pathway. This means, the body cannot deliver the stimuli to these muscles in the natural way as because the neural pathway is damaged due to spinal cord injury, vascular accident, other neural disorders and likewise and that is why another method (FES) is required for delivering these stimuli without using the nervous system (Peckham, P. H. et al., 2005). This method is widely used for providing neural rehabilitation to patients (Granat, M. H.et al., 1993 and Chantraine, A. et al., 1999). The need of FES is discussed in detail in the next section.

NEED OF FES

FES has numerous applications in the medical field. Providing neural rehabilitation to patients is undoubtedly the most important application of FES (Peckham, P. H. et al., 1992). Loss or impairment in the ability of muscle movement or sensation is called Paralysis which is caused by disruption of communication of nerve impulses along the pathway from the brain to the muscles. The primary causes of paralysis are spinal cord injury, vascular accidents, strokes or other existing disorders of nervous system (Birbaumer, N. et al., 2008). FES is manly used for improving the overall muscle contraction, assisting the patients in walking, reinstating the upper extremity and lower extremity to former state, increasing dexterity and strength in the limbs, restoring function in bladder, bowel, respiratory system and such (Stephenie Anna Haapala et al., 2008 and Chan, C. K., 2008). Apart from rehabilitation, FES has found application in several other domains like reduction of pain, improvement in blood circulation, inhibition

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of urinary bladder, influencing auditory and visual systems, reduction of seizures caused by epilepsy, excitation of spinal cord neural networks, increase in bone growth, prevention of scoliosis and many more (Alojz R. Kralj et al.,2000). FES basically helps to provide a better quality of life to the users.

Electromyography (EMG)

A myoelectric signal or a motor action potential is an electrical impulse which is responsible for producing contraction of muscle fibers in human body. These signals travels between muscles and nervous system thus regulating the movement of different parts of the body (Kumar, S.et al., 1996). The technique concerning the development, recording and analysis of these myoelectric signal is called Electromyography (EMG). Whenever the connection between brain and the muscles i.e. the neural pathway between them gets affected, the myoelectric signals generated within the muscles also change. As a result, study of these electrical signals generated is a very useful tool for diagnosing any abnormality in the muscles and also in the entire motor system. It was primarily used as a diagnostic tool earlier but with the advent of technology, the applications of EMG increased substantially (Reaz, M. B. I. et al., 2006 and 42 Merletti, R. et al., 2004). One of the most important applications of EMG is obtaining control of limbs artificially which is done mainly by using it to control FES and providing rehabilitation to the patients.

PURPOSE AND USEFULNESS OF EMG BASED FES

EMG based FES serves a very versatile platter of purpose and usefulness. To begin with, we must on the very onset mention about its role in treating muscular diseases and modeling various muscular models for the same purpose. EMG based FES has been used to construct muscle model based on the relationship between the frequency of stimulus and force. Another kind of physiological muscle model was also built based on the relationship of EMG and force. Many similar kinds of EMG-force relationship based (Perry, J. et al., 1981) experiments were carried out by many researchers across the globe to establish proper estimation required to build such models. To study muscle tone EMG based FES has been used with the help of muscle sound application. Muscle torque relationship using artificial neural network was also establish using EMG based FES (Hincapie, J. G. et al., 2009). Not only in the field of voluntary muscles but EMG based FES has been also applied to treat denervated muscle and smooth muscles present in bladder walls. Another very common area of application is fatigue compensation of different kinds of muscles present in human body.

The second most common area of application is for treating paralytic patients. Various degrees and extent of paralysis has been successfully treated using EMG based FES. Studies revealed that EMG based FES method was successful in making a paraplegic patient, stand up on his own feet. It is also many a times used to train extremity coordination in hemiplegic patients. Researchers and scientists had also developed artificial neural network based patient responsive ambulation aids for paraplegic conditions using EMG based FES. (C Komiyama et al, 1994)(D Popovic et al, 2007)

Several control strategies for limb extremity movements as well as joint models are designed using EMG based FES.(C A Kirkwood et al, 1989) There are instances where special EMG based FES models have been developed hand function restoration with patient driven loop. This technique is non-invasive and hence has a very promising field of application. Control strategies for orthotic exoskeleton of hand

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

are also designed with EMG based FES. For both upper and lower limb joints various kinds of musculoskeletal models have been developed. Such models are most prevalent for elbow joints or 2 joint systems of lower limbs. (M hayashibe et al, 2009)

In the continuation with the above field of application it is important to mention another common area of EMG based FES system application which is gait analysis. Many dynamic musculoskeletal models have been developed since long time for gait application. EMG based FES has been used to treat drop foot correction by studying the effect of walking speed changes on tibialis as obtained from its gait analysis. in people having partial paralysis gait initiation has been applied with the help of electromyography triggered FES. (A Dutta et al, 2009)

Electromyography based functional electrical stimulation can also be used for the purpose of neuromuscular rehabilitation. Literature suggests the Neuromuscular Modeling with Full Physiological Dynamics has been built using this methodology. Even FES evoked leg exercise has been designed for patient suffering from spinal cord injuries with the help of electromyography. (K Subramanya et al) (R.E.S. Shalaby, 2011)

Post stroke rehabilitation also needs special mention amongst the different areas of application of EMG based FES. Different studies are there that presents event related desynchronization modulated FES for post stroke rehab purpose using EMG. Neuromuscular stimulation are given after stroke to patients using this method. Also researchers used EMG based FES to aid walking in stroke patients.(M takahashi et al, 2012)

EMG BASED FES SYSTEMS

Since Nineties the usefulness of EMG based FES has been recognized and harnessed by researchers and scientists for the purpose of rehabilitation. It is found to serve as treatment for a wide range of diseases and disabilities. In this section a brief overview on the application areas of EMG based FES that has been in research or practice since 1990 is given. It will mainly cover the review about the rehabilitation of paralytic patients, stroke patients, upper and lower limb rehab, functioning of denervated musculatures, gait analysis and neurorehabillitation.

Paralysis

Chifuyu Komiyama *et al* long back in 1994 applied the idea of using EMG based FES for making paraplegic patients stand up. They developed an open loop control mechanism as such in which EMG was recorded from the quadriceps femoris and gluteus maximus muscles using surface electrodes, at a stand up posture of non paralyzed persons based on which they designed the stimulus that is to be applied to the paraplegic patient. The functional electrical stimulus was a constant current signal having amplitude modulated monophasic pulses. They did set a standard FES pattern for all paralyzed patients and then customized those stimulus patterns according to the requirement of different patients.

Daniel Graupe *et al* applied neural network theory to implement FES based on EMG on patients suffering from paraplegia. They used an ART-1 neural network scheme for the purpose of controlling the functional electrical stimulus that would facilitate patient responsive ambulation by paraplegic patients. They proposed such an ART-1 NN that would discriminate above lesion upper trunk EMG time series for activating stand up and walking functions and also control the electrical stimulus levels using

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

response EMG signals. Thus they successfully designed a single system effective for two purposed i.e. pattern recognition and control. Their system was so designed that it also gave the patients, space for manual over ride in case any error occurs.

In 2011 CH Chou *et al* implemented the idea of EMG based FES system in order to train hemiplegic patients for extremity coordination. It so happens in hemiplegic condition that not only does drop foot occurs but the upper extremity limits too while the patient tries to walk. In general only FES is applied to the drop foot to cure but this lacks the upper limb coordination and hence the researchers implemented the integration of a two channel EMG to improve the bad walking posture of CVA patients and also prevent their frequent falling down with enabling proper coordination between the drop foot and the upper limb swings. This not only helped in treating such cases of CVA patients but also corrected their bad standing or walking postures.

Gait

Anirban Dutta *et al* in 2009 proposed a gait model for partially paralytic people based on EMG triggered functional electrical stimulus. In people who suffered from incomplete spinal cord injury (iSCI) the muscle activities can be partially regained with FES triggered by electromyogram based gait event detector i.e. an EMG trigger. It emerges as a more natural command source rather than manual switches and hence provide improved coordination during gait.

C A Bryne *et al* in 2006 conducted an research study that aimed at producing an optimized stimulation envelope for drop foot correction. They employed EMG pattern of tibialis anterior obtained during healthy gait to design the required stimuli emvelope. Their objective was not to alter the walking speed of the person but was to provide data to enable the adaptation of the envelope with the walking speed. Their work result suggested that the FES envelope that produces larger EMG amplitude for toe off burst than the heel strike burst, would be more appropriate for hemiplegic patient's walking speed.

Muscle

Takashi Watanabe *et al* in 1999 developed a simplified model for electrically stimulated muscle based on EMG signals with a force frequency relationship for the application of FES. The EMG signals were obtained from neurologically intact persons. The relationship between force and frequency of electrically stimulated fast and slow type muscles was predicted in their muscle model. This model also described the summation of muscle contraction. All this finding were suggested to decrease the parameter identification time for patients on clinical purpose.

In 1994, T Yamamoto *et al* applied the idea of EMG along with muscle sound to assess muscle tonus, muscular strength, its tolerance and fatigue level of electrically stimulated muscles for the purpose of FES. This study established a strong relationship between muscle sound and inner pressure of fascia and also proved that muscle sound is a very good parameter to determine muscle tonus.

Mitshuhori Hayashibe et al in 2009 proposed an EMG to force estimation for voluntary muscle contraction that has a manifold application in human machine interface, rehabilitation and robotics. EMG based FES model of activation patterns of separate individuals can be very fruitful for estimating of muscle force. Previous to their study only macroscopic physiology of "hill type model" was prevalent

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

but with this study they successfully established a microscopic physiology for the same as well. This model brought about the understanding of internal biophysical dynamics and also revealed new insights about neuromuscular activations.

Zhan Li *et a*l in 2012 found that it is important to have torque estimation and prediction as feedback information for actual muscle state in order to make a good adaptive FES control. For this purpose they chose evoked EMG signals as according to them it was very correlated with FES induced torque under variety of muscle conditions. Hence in order to establish a repation between this evoked EMG and FES muscle torque they applied a nonlinear ARX model or NARX model of recurrent neural network. This RNN model identified and predicted FES induced muscle dynamics with Eemg. This is a very robust model and gives promising results.

In 2002 Winfired Mayr *et al* in one of their paper presented the fact that EMG based FES system can be very useful to regenerate, restore and maintain denerveted muscles. To effectively activate such denerveted musculature using surface EMG electrodes long duration biphasic impulses are required habing a pulse width varying from 1- to 300 ms with amplitude ranging from +- 100 to -+250 mA. But such situation may arise safety issues for application purpose. Though no such clinical set up has yet been developed yet many experimental purpose protypes are made which aims at addressing this safety issues at the same time serve the purpose.

Jeremey Laforet and David Guiraud developed a model for smooth muscle for FES application based on EMG in 2008. Their model enabled the stimulation of FES induced contraction, their smooth muscle model for bladder wall included muscle cell calcium dynamics. Their work was unique in a sense that they used this EMG based FES system in order to establish relation between calcium concentration and electric potential of cell membrane. They also refined the mechanical model and validated their results using simulation models. Hence not only to determine muscle fatigue or activate muscles, but EMG based FES system is also found to be used for biochemical studies.

Limb Control

In year 1995 William K Durfee *et al* carried an experimental verification of a lower limb 2 joint model with EMG based FES for gait application. Their model had 10 links, 14 degrees of freedom and aimed at facilitating the controller design for maximized artificial gait system performance. Primarily they conducted experiments to know whether a subset of the full model can be used for force and motion prediction of actual limbs in order to ensure high level of accuracy. The angle of knee and hip joints were measured for passive initial condition releases and also for activating single muscles by FES.

Musculoskeletal Joint models stared developing way back in nineties on the principle of EMG based FES. Takashi Watanabe *et al* in 1998 conducted experimental tests on such a model for FES applications. They tested this model on a normal subject with surface stimulus . the stimulation data were based on EMG signals as obtained from his upper limb muscles. Parameters used in this model was based on force measured under isometric condition. They found that the output for the tuned model matched well with that of the experiment of a single muscular stimulation at high amplitude.

Han chang in 2000 developed a versatile multichannel direct synthesized electrical stimulator for the application of FES. This stimulator developed by them was an arbitrary waveform stimulator. The authors proposed the idea of an element envelope and applied digital signal processor for synthesizing the stimulation patterns having high resolution. Their system aimed at reducing memory requirements along with keeping the space for flexibility of waveform type. The system was also well equipped to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

adjust its stimulation parameters dynamically. The output stage of their system provided a biphasic voltage controlled constant current patterns that had wide enough bandwidth and high compliance. Thus this EMG based FES system had the capability of feedback processing and also the flexibility for pattern generation.

In 2004, Ying Han Chiou *et al* proposed a non invasive scheme of FES based on patient driven control loop for restoring hand movement functions. It was so designed that any patient could use their well-to-do limb extremities to control the functional electrical stimulus that would in turn excite or activate their paralyzed muscle cells. The system would receive EMG signal as trigger to serve the purpose of system initialization and adjustment of electrical stimulation parameter. Thus a complete patient driven control loop was developed by the researchers with an EMG based FES system.

Mathew Dicicco *et al* in 2004 constructed a very low profile light weight orthotic exoskeleton to restore the dexterity of upper limb extremity that would be controlled by patient residual EMG signal. They compared several such systems of EMG based FES for this restoration purpose and found that electromyogram signals retrieved from ipsilateral biceps is supposedly the best option for developing a very reliable natural reaching and pinching algorithm.

Neuro-Muscular Rehabilitation

In 2007 R Lyne *et al* in one of their review paper presented different comprehensive uses of EMG based FES for neuro-rehabilitation. The main functional application in this field is to active paralyzed neuro-musculature to get tasks done properly with high precision. It is mentioned in this study that apart from post stroke rehab other common neuronal support that can be provided using FES based on EMG is renervation of respiratory tract, urinary bladder, prevention of neuromuscular collapse and prophylaxis of deep venous thrombosis and similar kinds.

Stroke

R Li *et al* developed a novel type of EMG based FES robotic system that aimed at post stroke rehabilitation for wrist movement impairment in the year 2008. The control algorithm of this robot was so designed that it interacted with the voluntary motor inputs from the stroke patient. They represented their result by root mean square error and root mean squared jerk of the trajectories of wrist tracking as obtained from their tasks that were supported by FES. It showed that the integrated robot system was effective in improving tracking performance for cases of both acute and chronic stroke patients.

Juan Nicholas Cuenca *et al* published in a paper in 2004 that how EMG based FES system helps in physiotherapy after a person had a stroke. It is said to be a very effective tool for the purpose of maintaining the trophism of muscle. It also helps in reducing the appearance of edema and fibrosis in paretic muscles that suffered post effects of stroke. If combined with cognitive activation this system would also allow the voluntary access of patient to the concerned muscles.

In 2003 Yukihiro Hara presented various area of application of rehabilitation with EMG based FES for stroke patients. In many cases after a stroke it so happens upper limbs becomes hemiparetic. EMG triggered FES can in particular improve the motor function of such limbs. It has been found that triggered FES is much more effective than non triggered one as it induces greater muscle contraction that is proportional to integrated voluntary EMG signals. Wrist movement, shoulder flexion, finger extension

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

can be improved in stroke patient if a daily basis therapy of EMG based FES is applied. Also sensory motor impairment can be treated with the same as the cerebral blood flow increases in sensory motor cortex of brain when EMG triggered FES is applied.

In a review article published in 2014 by "Different strokes" organization of United Kingdom the beneficial aspects of EMG based FES system for walking aid purpose for stroke patients has been presented elaborately. People often suffer from drop foot syndrome after stroke. This can be cured applying EMG based FES. It brings about natural walking pattern in the patient along with the increase in the walking speed too. They are even able to walk in uneven surfaces. People become enabled for walking relatively longer distances. The frequent tendency to fall down also reduces manifold as it reduces spasticity. Regular use of this FES triggered by EMG system actually treats the stroke survivor pretty well since he becomes able to walk with appreciable speed after a certain period of therapy time all by himself when the FES is switched off. A brief summary of the above discussion is given in Table 1.

Table 1. Chronological overview of EMG based FES applications

Year of Publication	Author	Category	Objective	Area of Improvement
1994	Chifuyu Komiyama *et al*	Paralysis	To Make a Hybrid FES orthosis to make patient stand up in paraplegic condition	Experiments were conducted on healthy subjects
1994	D Graupe, H Kordylewski	Paralysis	To use Neural network to control FES in paraplegics for patient-responsive ambulation	
1994	T Yamamoto, H Takano	Muscle Activation	To study the Application of muscle sound to assess muscle tonus of the electrically stimulated muscle for FES	Not applicable for fully paralytic musculature
1995	D Graupe, H Kordylewski	Paralysis	To use Artificial neural network for controlling FES for patient responsive ambulation in paraplegic condition	
1998	Takashi Watanabe *et al*	Joint Dysfunction	To conduct Experimental tests of a musculoskeletal model of the elbow joint for FES applications	The reproducibility of the model was not satisfactory
1999	Takashi watanabe *et al*	Muscle Modelling	To present An approach of muscle model with a stimulus frequency-force relationship for FES applications	The muscle model needs some improvisation for obtaining more accurate predictions
2000	Han-Chang Wu, Shuenn-Tsong Younz, Te-Son Kuo	Limb Dysfunction	To make a versatile multichannel Direct-Synthesized Electrical Stimulator for Applications of FES	
2004	M DiCicco, L Lucas, Y Matsuoka	Limb Dysfunction	To compare the control strategies for an EMG controlled orthotic exoskeleton for the hand	
2004	Ying-Han Chiou, Shih-Ching Chen, Jin-Shin Lai, Te-Son Kuo	Limb Dysfunction	To make a Noninvasive Functional Electrical Stimulation System with Patient-Driven Loop for the purpose of Hand Function Restoration	The system is more time-consuming in comparison to other systems
2007	L R Sheffler, J Chae	Neuromuscular Rehabilitation	To study the application of Neuromuscular electrical stimulation in neurorehabilitation	

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Continued

Year of Publication	Author	Category	Objective	Area of Improvement
2007	C A Byrne, D T O'keeffe, A E Donnelly, G M Lyons	Gait	To study the Effect of walking speed change on tibialis anterior EMG during healthy gait for FES envelope design in drop foot correction	Only male subjects were considered for the study
2008	R Li, X L Hu, K Y Tong	Stroke	To develop a Combined EMG-driven system FES for post-stroke rehabilitation	Smooth tracking of the FES system was hindered because it depended on the subject
2008	E Lazar, J Nicolás Cuenca	Stroke	To study the effectiveness of Functional electrical simulation in treating stroke patients.	
2009	A Dutta, R Kobetic, R J Triolo	Gait	To study gait with EMG triggered electrical stimulation in people having partial paralysis	
2009	M Hayashibe, D Guiraud, P Poignet	Muscle Modelling	To make EMG-to-force estimation with full-scale physiology based muscle model	More accurate estimation of muscle force from EMG signal is desired
2011	C Chou, S Chen, Y Hwang, C Ho, C Chen	Paralysis	To study the application of FES for hemiplegia in order to train extremity coordination	
2013	Yukihiro Hara	Stroke	To use Functional Electrical Stimulation in Stroke Patients for rehabilitation purpose	
2014	"Different strokes" organization	Stroke	To do drop foot correction in patients who suffered from stroke.	

OTHER METHODS OF FES

It is quite evident from the above discussion that the most popular method of controlling FES systems is by using EMG signals. Apart from EMG signals, other signals like ENG and EEG have also been used for driving these kind of systems(Thomas Sinkjaer et al.,2003). EEG or Brain Computer Interface based FES systems were first introduced in late 1990s(Lauer RT, 1999 and Stein RB,1980). A number of research work has been undertaken for developing stimulators for controlling hands, limbs and other parts of the body(An H Do et al., 2012, Ang, K. K.et al., 2013). Besides EEG, researchers have tried to utilize another biosignal namely Electroneurography(ENG) signal. Both EEG and ENG signal has been used as controlling signal in laboratory but ENG signals have found application mostly as feedback signal for controlling any device.

After EMG signals, EEG signals are the most preferred signals to be used as command signals for stimulators. Over the past few years, scientists have explored this particular area quite intensively, specially how this technology can be utilized for the improvement of the quality of life of people suffering from different neural disorders. In 2005, a study was performed by Gert Pfurtscheller et al. to use an asynchronous EEG based Brain Computer Interface system for controlling an FES stimulator. The stimulator was designed to assist a tetraplegic patient for improvement of his hand grasp without any help from outside sources. The subject was required to imagine foot movement and the brain wave

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

generated by the patient while doing so was recorded by the EEG machine and converted to command signals by the BCI system to control the Functional Electrical stimulator. The result obtained by them was quite satisfactory because the patient was successful in grasping the desired object with his paralyzed hand completely on his own.

Another study was conducted in 2008 to provide rehabilitation to stroke patients. A BCI based FES system was proposed in the study in which the motor intention of subjects were recorded in form of EEG signals. Basically they tried to extract the Event-Related Desynchronization from which the motor intention of the subjects could be detected and these signals were then used for improving the limb activity. The main objective of using FES in this system was to stimulate the relevant muscles, which generates a sensory feedback to the brain, creating kind of a pseudo loop between the brain and the muscles.

FES DEVICES

The FES based systems are basically of two types external devices and implants. Both of these systems have some advantages and disadvantages associated with them. The FES system based implants or devices have evolved to a large extent in the past few decades. The development of the implants has been quite rapid and so far three generations of implants can be properly distinguished. The first generation of implants mainly consisted of simple electronic circuits and no measures were taken to eliminate the noises incorporated with the signal by displacement of the receiver with respect to the transmitter (Syms 3rd, CA et al., 1997 and Brindley GS.et al., 1994). In the next generation of FES implants, the power was derived from an inductive link before but the only difference was that the system was modified to provide stabilized output. The difference between the second and the third generations of stimulators is that, the third generation devices were powered by a battery instead of an inductive link. Vienna FES implant is one of the earliest designs used to assist tetraplegic patients. It was implanted for the first time in 1982(Winfried Mayr et al., 2001). It is basically an 8 channel implant and powered by inductive link. This design has been modified and few different designs of implants were derived from it which are still at the research level.

In 2004 (Fumihiro Sato et al.), A new contactless transmission based design was proposed. This system consisted of two parts. One coil, placed outside the limb and another part implanted within the body. This implanted part consists of two segments of coil- one to receive the signal and another to transmit the same. The RF frequency required for both reception and transmission of signals are different. In this design both the power and the signal required for stimulating the muscles get transmitted together from outside but they get separated inside the muscles. They also introduced a capacitor in the secondary coil to increase the power output of the entire system.

The development process of the FES system had several hindrances among which the problem of proper muscle recruitment is of utmost importance. Djordje Popovic et al. in 2005 evaluated the recruitment properties of the BION implant. This implant is basically a compact and self sufficient implant which can be inserted within the muscle. This implant is quite suitable for eliminating the problem of muscle recruitment (Djordje Popovic, 2007).

In another case (Anirban Dutta, 2009), the FES system was used to assist patients suffering from incomplete spinal cord injury while walking. In this system, both the EMG data acquisition and stimulation was done by means of surgically implanted electrodes. The anatomical positions suitable for placing these electrodes were detected prior to this implantation with the help of surface EMG electrodes. The stimulation was given to the muscle through this implanted stimulator.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Coming to the external devices, different forms of these devices have also been utilized for the remedy of neuromuscular diseases or neurological disorders. A number of research work has been performed in order to provide remedy to such problems. One of the examples of this kind of devices is exerciser. The first exerciser was an FES based cycling system designed by Petrofsky and Phillips (J.Petrofsky et al, 1984). This system was mainly suitable for people suffering from Spinal Cord Injury(SCI). In 2009, a Leg Stepping Trainer was developed to provide rehabilitation to these particular category of patients. They are mostly incapable of doing exercise because of their paralyzed leg. This isokinetic device maintained a static speed throughout the period of exercise and mainly strengthened the weaker muscles and restricted any extra movement or extra stepping effort of the leg (Nur Azah Hamzaid et al., 2009). It was actually used efficiently for two reasons viz. leg resistance training and cardiorespiratory training at comparatively higher speed.

Yoon-Kyu Song et al.(2007) proposed a new fiber optic based design of an FES device. In this device, light energy was converted to electrical energy which in turn was used to stimulate a sciatic nerve of a frog. At first, infrared pulses were made to enter a multimodal optical fiber, which then guided these pulses to a photovoltaic device. At this photovoltaic cell the light got converted to current and this current was responsible for stimulating the target nerve. This approach might be quite useful in fabricating new fiber optics based neuromotor prostheses.

The timing for application of the electrical pulses in any FES system is very crucial to get the desired effect and should be taken into consideration at each and every instant. An wearable device was developed in 2009 by Goro Obinata et al. for detecting the state when the switching of the walking motion is taking place. They also used Artificial Neural Network(ANN) for estimating this phase switching time. The results obtained by them showed that this wearable device rendered quite a good estimation of this timing.

Over the years, huge number of studies have been performed for modification of the FES systems. One of the aspects on which the scholars have concentrated is to increase the flexibility of the electrical pulse delivery system. Haibin Wang et al.(2012)devised an FES system which was capable for regulating the amplitude, frequency and waveform of the electrical pulses which are to be applied to the muscles. This system has application both the domains of medicine and research.

FES systems have been used even for rehabilitation of knee joint. The FEXO knee (Yong Ren et al., 2014) is an example of such a device. In this device, the stimulator was combined with an exoskeleton to make the entire system more effective. A number of similar devices have been designed, tested and applied clinically. Researchers are putting an enormous amount of effort for modifying these devices or systems so as to devise better and more sophisticated systems which will help people suffering from the previously mentioned maladies regain their former state of life.

FUTURE RESEARCH DIRECTIONS

Neuromuscular diseases (nmds) are listed to be one of those maladies that pose severe threat on the human race. However, the only feasible option to reinstate the patients, suffering from these genre of disability, to their former state is to provide them with next generation rehabilitation. These rehabilitative aids not only help them to overcome their disabilities but also gift them a better quality of life.

Previously, people victimized with such diseases could not have dreamt of having their life restored back to their healthy state. But with the advent of technology, such as EMG based FES, researchers and scientists have been successful to quite a good extent to realise their dreams which were once next

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to impossible. However, there are many loop holes that are yet to be filled in order to make the next generation of rehabilitative devices flawlessly effective.

One of such major drawbacks in present day's EMG based Functional Electrical Stimulators is that, most of these devices are implantable ones which demands the patient's body to be operated, that causes pain and hassle for the patient. Besides, there can also be biocompatibility issues that can result in infection of the patient concerned or in worst case rejection of the device. Also, those devices which are placed percutaneously, can cause discomfort and swelling.

As because the entire method relies on the application of electrical pulses to the human body, hence there always remains a probability for the patient to get affected by electric shock. This can often result in worsening of the affected muscles or even dysfunctioning of healthy musculature. Again if the duration or the amplitude of the applied pulses exceed the permissible limit, it can impose burns or other adverse effects.

The above mentioned shortcomings of EMG based FES systems can actually form the basis for the future direction of research works. For example, power isolation can be provided between the device and the patient so as to avoid mishap caused by electric shock. Apart from that, optimization of power consumption of these devices can also increase their acceptability.

CONCLUSION

The main aim of this chapter is to put up the technique of EMG based FES as a promising rehabilitation method for patients with neuromuscular impairment. From time immortal neuromuscular maladies have been a great reason of concern for human race to which the most suitable remedy is functional electrical stimulation. FES is mainly used to provide short duration electrical pulses to those muscles which suffer from nervous disorder in order to treat them. The EMG signals from the affected muscular region is used as biofeedback to the functional electrical stimulator Though from quiet a long time it has been in use for many clinical application but still there are many aspects of this method that is yet to be explored. Various categories of disorders can be treated using this system of which the most popular area of treatment are paralytic patients having hemiplegic, paraplegic or tetraplegic conditions, stroke patients and post stroke neuromuscular rehabilitation, patients suffering from muscular fatigue and limb movement control for physically disabled patients. Apart from these function electrical stimulation using EMG is also widely used for the purpose of Gait analysis of healthy as well as physically handicapped persons. But despite of having a lot of advantages like all other systems it too poses some certain disadvantages like risk of electrical shock, bio-incompatibility issues (in case of implantable ones), and relatively high power consumption. From quiet a long time this method has found a prominent usage in treating many patients but yet till now there is a lot of scope for improving this entire EMG based FES system in order to make it full proof and robust.

REFERENCES

Ang, K. K., & Guan, C. (2013). Brain-Computer Interface in Stroke Rehabilitation. *Journal for Corrosion Science and Engineering*, *7*(2), 139–146. PMID:25120465

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Baldi, J. C., Jackson, R. D., Moraille, R., & Mysiw, W. J. (1998). Muscle atrophy is prevented in patients with acute spinal cord injury using functional electrical stimulation. *Spinal Cord*, *36*(7), 463–469. doi:10.1038j.sc.3100679 PMID:9670381

Birbaumer, N., Murguialday, A. R., & Cohen, L. (2008). Brain-computer interface in paralysis. *Current Opinion in Neurology*, *21*(6), 634–638. doi:10.1097/WCO.0b013e328315ee2d PMID:18989104

Brindley, G. S. (1994). The first 500 patients with sacral anterior root stimulator implants: General description. *Spinal Cord*, *32*(12), 795–805. doi:10.1038c.1994.126 PMID:7708419

Byrne, C. A., O'keeffe, D. T., Donnelly, A. E., & Lyons, G. M. (2007). Effect of walking speed changes on tibialis anterior EMG during healthy gait for FES envelope design in drop foot correction. *Journal of Electromyography and Kinesiology*, *17*(5), 605–616. doi:10.1016/j.jelekin.2006.07.008 PMID:16990012

Chan, C. K. (2008). A preliminary study of functional electrical stimulation in upper limb rehabilitation after stroke: An evidence-based review. *Hong Kong Journal of Occupational Therapy*, *18*(2), 52–58. doi:10.1016/S1569-1861(09)70003-8

Chantraine, A., Baribeault, A., Uebelhart, D., & Gremion, G. (1999). Shoulder pain and dysfunction in hemiplegia: Effects of functional electrical stimulation. *Archives of Physical Medicine and Rehabilitation*, *80*(3), 328–331. doi:10.1016/S0003-9993(99)90146-6 PMID:10084443

Chesler, N. C., & Durfee, W. K. (1997). Surface EMG as a fatigue indicator during FES-induced isometric muscle contractions. *Journal of Electromyography and Kinesiology*, *7*(1), 27–37. doi:10.1016/S1050-6411(96)00016-8 PMID:20719689

Chiou, Y. H., Chen, S. C., Lai, J. S., & Kuo, T. S. (2004). A noninvasive functional electrical stimulation system with patient-driven loop for hand function restoration. In *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (Vol. 2, pp. 4130-4132). 10.1109/IEMBS.2004.1404152

Chou, C., Chen, S., Hwang, Y., Ho, C., Chen, C., Chen, S., & Chen, Y. (2011). Application of FES for hemiplegia in extremity coordination training. In *5th International Conference on Bioinformatics and Biomedical Engineering* (pp. 1-4).

Dariush, B., Parnianpour, M., & Hemami, H. (1998). Stability and a control strategy of a multilink musculoskeletal model with applications. *IEEE Transactions on Biomedical Engineering, 45*(1), 3-14.

Demosthenous, A., Jiang, D., Pachnis, I., Liu, X., Rahal, M., & Donaldson, N. (2008). A programmable ENG amplifier with passive EMG neutralization for FES applications. In *IEEE International Symposium on Circuits and Systems*, (pp. 1552-1555). 10.1109/ISCAS.2008.4541727

DiCicco, M., Lucas, L., & Matsuoka, Y. (2004). Comparison of control strategies for an EMG controlled orthotic exoskeleton for the hand. In *Proceedings of International Conference on Robotics and Automation*, (Vol. 2, pp. 1622-1627). 10.1109/ROBOT.2004.1308056

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Do, A. H., Wang, P. T., King, C. E., Schombs, A., Cramer, S. C., & Nenadic, Z. (2012). Brain-computer interface controlled functional electrical stimulation device for foot drop due to stroke. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 6414-6417). 10.1109/EMBC.2012.6347462

Durfee, W. K., Chiou, J. T., & Beck, H. L. (1992). Dynamic musculo-skeletal model for FES-aided gait applications. In *14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (Vol. 4, pp. 1335-1336). 10.1109/IEMBS.1992.5761817

Durfee, W. K., & Dennerlein, J. T. (1989). EMG as a feedback signal in surfaces FES applications: issues and preliminary results. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 1009-1010). 10.1109/IEMBS.1989.96059

Durfee, W. K., & Dunn, B. N. (1995). Experimental verification of a 2-joint, lower limb musculoskeletal model for FES-aided gait applications. In *IEEE 17th Annual Conference Engineering in Medicine and Biology Society,* (Vol. 2, pp. 1149-1150). IEEE.

Dutta, A., Kobetic, R., & Triolo, R. (2009). Development of an implanted intramuscular EMG-triggered FES system for ambulation after incomplete spinal cord injury. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 6793-6797). 10.1109/IEMBS.2009.5333980

Dutta, A., Kobetic, R., & Triolo, R. J. (2009). Gait initiation with electromyographically triggered electrical stimulation in people with partial paralysis. *Journal of Biomechanical Engineering*, *131*(8), 081002. doi:10.1115/1.3086356 PMID:19604014

Ebashi, S., Endo, M., & Ohtsuki, I. (1969). Control of muscle contraction. *Quarterly Reviews of Biophysics*, *2*(04), 351–384. doi:10.1017/S0033583500001190 PMID:4935801

Fye, W. B. (1994). A history of the origin, evolution, and impact of electrocardiography. *The American Journal of Cardiology*, *73*(13), 937–949. doi:10.1016/0002-9149(94)90135-X PMID:8184849

Granat, M. H., Ferguson, A. C. B., Andrews, B. J., & Delargy, M. (1993). The role of functional electrical stimulation in the rehabilitation of patients with incomplete spinal cord injury-observed benefits during gait studies. *Spinal Cord*, *31*(4), 207–215. doi:10.1038c.1993.39 PMID:8493035

Graupe, D., & Kordylewski, H. (1994). Neural network control of FES in paraplegics for patient-responsive ambulation. In *IEEE International Symposium on Circuits and Systems*, (Vol. 6, pp. 249-252). 10.1109/ISCAS.1994.409574

Graupe, D., & Kordylewski, H. (1995). Artificial neural network control of FES in paraplegics for patient responsive ambulation. *Biomedical Engineering. IEEE Transactions on*, *42*(7), 699–707.

Graupe, D., Kralj, A., Basseas, S., & Kohn, K. H. (1982). EMG parameter identification for controlling electrical stimulation of peripheral nerves to provide certain paraplegics with primitive walking functions. In *21st IEEE Conference on Decision and Control*, (Vol. 21, pp. 345-350). 10.1109/CDC.1982.268457

Haapala, S. A. (2008). *Joint biomechanics during progressive resistance and submaximal functional electrical stimulation-leg cycle ergometry in spinal cord injured people*. ProQuest.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hamzaid, N. A., & Davis, G. M. (2009). Health and Fitness Benefits of Functional Electrical Stimulation-Evoked Leg Exercise for Spinal Cord–Injured Individuals. *Topics in Spinal Cord Injury Rehabilitation*, *14*(4), 88–121. doi:10.1310ci1404-88

Hamzaid, N. A., Fornusek, C., Ruys, A. J., & Davis, G. M. (2009). Development of an isokinetic FES leg stepping trainer (iFES-LST) for individuals with neurological disability. In *IEEE International Conference on Rehabilitation Robotics*, (pp. 480-485). 10.1109/ICORR.2009.5209568

Hara, Y. (2013). Rehabilitation with Functional Electrical Stimulation in Stroke Patients. *Int J Phys Med Rehabil, 1*(147).

Hayashibe, M., Guiraud, D., & Poignet, P. (2009). EMG-based neuromuscular modeling with full physiological dynamics and its comparison with modified hill model. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 6530-6533). 10.1109/IEMBS.2009.5333147

Hayashibe, M., Guiraud, D., & Poignet, P. (2009). EMG-to-force estimation with full-scale physiology based muscle model. In *International Conference on Intelligent Robots and Systems*, (pp. 1621-1626). 10.1109/IROS.2009.5354644

Hincapie, J. G., & Kirsch, R. F. (2009). Feasibility of EMG-based neural network controller for an upper extremity neuroprosthesis. *Neural Systems and Rehabilitation Engineering*. *IEEE Transactions on*, *17*(1), 80–90.

IJzerman, M. J., Renzenbrink, G. J., & Geurts, A. C. (2009). *Neuromuscular stimulation after stroke: from technology to clinical deployment*.

Johnston, D., & Wu, S. M. S. (1995). *Foundations of cellular neurophysiology*. Cambridge, MA: MIT Press.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (Eds.). (2000). *Principles of neural science* (Vol. 4). New York: McGraw-Hill.

Kang, X. Y., Liu, J. Q., Tian, H. C., Du, J. C., Yang, B., Zhu, H. Y., ... Yang, C. S. (2014). Fabrication and degradation characteristic of sputtered iridium oxide neural microelectrodes for FES application. In *27th International Conference on Micro Electro Mechanical Systems (MEMS)*, (pp. 616-619). 10.1109/MEMSYS.2014.6765716

Kaniusas, E., & Kaniusas, E. (2012). Fundamentals of Biosignals. Biomedical Signals and Sensors I: Linking Physiological Phenomena and Biosignals, 1-26.

Kirkwood, C. A., & Andrews, B. J. (1989). Finite state control of FES systems: Application of AI inductive learning techniques. In *Proceedings of the Annual International Conference of the IEEE Engineering in Engineering in Medicine and Biology Society, 1989. Images of the Twenty-First Century*. 10.1109/IEMBS.1989.96065

Komiyama, C., Saito, E., Tomita, Y., & Honda, S. (1994). Hybrid FES orthosis for standing up in paraplegia. In *Conference Proceedings of Instrumentation and Measurement Technology Conference*, (pp. 575-578). 10.1109/IMTC.1994.351894

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kralj, A. R., & Bajd, T. (1989). *Functional electrical stimulation: standing and walking after spinal cord injury*. CRC Press.

Kumar, S., & Mital, A. (Eds.). (1996). *Electromyography in ergonomics*. CRC Press.

Laforet, J., & Guiraud, D. (2008). Smooth Muscle Model for Functional Electrical Stimulation applications: Simulation of realistic bladder behavior under FES. In *Proceedings of 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 3702-3705). IEEE.

Lazar, E., & Nicolás Cuenca, J. (2008). *Functional electrical simulation (FES) in stroke*. Academic Press.

Li, R., Hu, X. L., & Tong, K. Y. (2008). Combined Electromyography (EMG)-driven system with functional electrical stimulation (FES) for poststroke rehabilitation. In *2nd International Conference on Biomedical Robotics and Biomechatronics, 2008. BioRob 2008.* 10.1109/BIOROB.2008.4762821

Li, Z., Hayashibe, M., Zhang, Q., & Guiraud, D. (2012). FES-induced muscular torque prediction with evoked EMG synthesized by NARX-type recurrent neural network. In *Proceedings of International Conference on Intelligent Robots and Systems*, (pp. 2198-2203).

Liberson, W. T., Holmquest, H. J., Scot, D., & Dow, M. (1961). Functional electrotherapy: Stimulation of the peroneal nerve synchronized with the swing phase of the gait of hemiplegic patients. *Archives of Physical Medicine and Rehabilitation*, *42*, 101–105. PMID:13761879

Mayr, W., Bijak, M., Rafolt, D., Sauermann, S., Unger, E., & Lanmüller, H. (2001). Basic design and construction of the Vienna FES implants: Existing solutions and prospects for new generations of implants. *Medical Engineering & Physics*, *23*(1), 53–60. doi:10.1016/S1350-4533(01)00014-5 PMID:11344008

Mayr, W., Hofer, C., Bijak, M., Rafolt, D., Unger, E., Sauermann, S., & Kern, H. (2002). Functional Electrical Stimulation (FES) of denervated muscles: Existing and prospective technological solutions. *Basic and Applied Myology*, *12*(6), 287–290.

Merletti, R., & Parker, P. A. (Eds.). (2004). *Electromyography: physiology, engineering, and non-invasive applications* (Vol. 11). John Wiley & Sons. doi:10.1002/0471678384

Miyake, Y. (2006). Electrooculography. Electrodiagnosis of Retinal Diseases.

Niedermeyer, E., & da Silva, F. L. (Eds.). (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

Obinata, G., Ogisu, T., Hase, K., Kim, Y., & Genda, E. (2009). State estimation of walking phase and functional electrical stimulation by wearable device. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 5901-5904). 10.1109/IEMBS.2009.5334827

Oh, S. J. (2003). *Clinical electromyography: nerve conduction studies*. Lippincott Williams & Wilkins.

Peckham, P. H., & Creasey, G. H. (1992). Neural prostheses: Clinical applications of functional electrical stimulation in spinal cord injury. *Spinal Cord*, *30*(2), 96–101. doi:10.1038c.1992.33 PMID:1589297

Peckham, P. H., & Knutson, J. S. (2005). Functional electrical stimulation for neuromuscular applications. *Annual Review of Biomedical Engineering*, *7*(1), 327–360. doi:10.1146/annurev.bioeng.6.040803.140103 PMID:16004574

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Perry, J., & Bekey, G. A. (1981). EMG-force relationships in skeletal muscle. *Critical Reviews in Biomedical Engineering*, *7*(1), 1–22. PMID:7042199

Petrofsky, J. S., & Phillips, C. A. (1984). The use of functional electrical stimulation for rehabilitation of spinal cord injured patients. *Central Nervous System Trauma*, *1*(1), 57–74. PMID:6400201

Pfurtscheller, G., Müller-Putz, G. R., Pfurtscheller, J., & Rupp, R. (2005). EEG-based asynchronous BCI controls functional electrical stimulation in a tetraplegic patient. *EURASIP Journal on Applied Signal Processing*, *2005*(19), 3152–3155. doi:10.1155/ASP.2005.3152

Popovic, D., Baker, L. L., & Loeb, G. E. (2007). Recruitment and comfort of BION implanted electrical stimulation: Implications for FES applications. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *15*(4), 577–586. doi:10.1109/TNSRE.2007.909816 PMID:18198716

Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biological Procedures Online*, *8*(1), 11–35. PMID:16799694

Ren, Y., & Zhang, D. (2014). FEXO Knee: A rehabilitation device for knee joint combining functional electrical stimulation with a compliant exoskeleton. In *Proceedings of International Conference on Biomedical Robotics and Biomechatronics*, (pp. 683-688). 10.1109/BIOROB.2014.6913857

Robinson, C. J., & Peckham, P. H. (1997). Guest Editorial FES Control Strategies, Methods of Selective Activation, Applications, and Test Models. *IEEE Transactions on Rehabilitation Engineering*, *5*(1), 1. doi:10.1109/TRE.1997.559343 PMID:9086388

Rushton, D. N. (1997). Functional electrical stimulation. *Physiological Measurement*, *18*(4), 241–275. doi:10.1088/0967-3334/18/4/001 PMID:9413861

Sato, F., Nomoto, T., Kano, G., Matsuki, H., & Sato, T. (2004). A new contactless power-signal transmission device for implanted functional electrical stimulation (FES). *IEEE Transactions on Magnetics*, *40*(4), 2964–2966. doi:10.1109/TMAG.2004.830416

Shalaby, R. E. S. (2011). *Development of an Electromyography Detection System for the Control of Functional Electrical Stimulation in Neurological Rehabilitation.* (Doctoral dissertation). TU Berlin.

Sheffler, L. R., & Chae, J. (2007). Neuromuscular electrical stimulation in neurorehabilitation. *Muscle & Nerve*, *35*(5), 562–590. doi:10.1002/mus.20758 PMID:17299744

Sinkjaer, T., Haugland, M., Inmann, A., Hansen, M., & Nielsen, K. D. (2003). Biopotentials as command and feedback signals in functional electrical stimulation systems. *Medical Engineering & Physics*, *25*(1), 29–40. doi:10.1016/S1350-4533(02)00178-9 PMID:12485784

Smorto, M. P., & Basmajian, J. V. (1974). Clinical electroneurography. *American Journal of Physical Medicine & Rehabilitation*, *53*(2).

Song, Y. K., Patterson, W. R., Bull, C. W., Zhang, J., Sheldon, C. R., Nurmikko, A. V., & Donoghue, J. P. (2007). Fiber Optic Guided Functional Electrical Stimulation with Microscale Photovoltaic Neurostimulator Devices. In *Conference on Lasers and Electro-Optics*. Optical Society of America. 10.1109/CLEO.2007.4452964

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Subramanya, K., & Kanakabettu, A. K. M. (2014). Functional electrical stimulation for neuro rehabilitation: A new design paradigm. *Short Communications Functional Electrical Stimulation for Neuro Rehabilitation. A New Design Paradigm*.

Syms, C. A. III, & House, W. F. (1997). Surgical rehabilitation of deafness. *Otolaryngologic Clinics of North America*, *30*(5), 777–782. PMID:9295253

Takahashi, M., Gouko, M., & Ito, K. (2008). Electroencephalogram (EEG) and Functional Electrical Stimulation (FES) System for Rehabilitation of Stroke Patients. In *Proceedings of 21st IEEE International Symposium on Computer-Based Medical Systems*, (pp. 53-58). 10.1109/CBMS.2008.45

Takahashi, M., Takeda, K., Otaka, Y., Osu, R., Hanakawa, T., Gouko, M., & Ito, K. (2012). Event related desynchronization-modulated functional electrical stimulation system for stroke rehabilitation: A feasibility study. *Journal of Neuroengineering and Rehabilitation*, *9*(1), 56. doi:10.1186/1743-0003-9-56 PMID:22897888

Wang, H., Guan, G., He, Q., Zeng, D., Leng, B., Xu, H., & Zheng, W. (2012). An electrical muscle simulator based on functional electrical stimulation. In *IEEE International Conference on Robotics and Biomimetics*, (pp. 1906-1911). 10.1109/ROBIO.2012.6491246

Watanabe, T., Futami, R., Hoshimiya, N., & Handa, Y. (1999). An approach to a muscle model with a stimulus frequency-force relationship for FES applications. *Rehabilitation Engineering*. *IEEE Transactions on*, *7*(1), 12–18.

Watanabe, T., Nozawa, T., Eom, G. M., Ohba, S., Futami, R., Hoshimiya, N., & Handa, Y. (1998). Experimental tests of a musculoskeletal model of the elbow joint for FES applications. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (Vol. 5, pp. 2594-2597). 10.1109/IEMBS.1998.744988

Winslow, J., Jacobs, P. L., & Tepavac, D. (2003). Fatigue compensation during FES using surface EMG. *Journal of Electromyography and Kinesiology*, *13*(6), 555–568. doi:10.1016/S1050-6411(03)00055-5 PMID:14573370

Wu, H. C., Young, S. T., & Kuo, T. S. (2002). A versatile multichannel direct-synthesized electrical stimulator for FES applications. *Instrumentation and Measurement*. *IEEE Transactions on*, *51*(1), 2–9.

Yamaguchi, H., Yu, W., Maruishi, M., Yokoi, H., Mano, Y., & Kakazu, Y. (2000). EMG automatic switch for FES control for hemiplegics using artificial neural network. In *Proceedings of the Sixth Conference on Intelligent Autonomous Systems*.

Yamamoto, T., & Takano, H. (1994). Application of muscle sound to assess muscle tonus of the electrically stimulated muscle for FES. In *Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Engineering Advances: New Opportunities for Biomedical Engineers.* 10.1109/IEMBS.1994.412011

This research was previously published in Biomedical Image Analysis and Mining Techniques for Improved Health Outcomes edited by Wahiba Ben Abdessalem Karâa and Nilanjan Dey; pages 337-355; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 45
A Novel Delay-Based GFSK Demodulator in 65 nm CMOS for Low Power Biomedical Applications

Meng Fu
University of Melbourne, Australia

Stan Skafidas
University of Melbourne, Australia

Iven Mareels
University of Melbourne, Australia

ABSTRACT

This article describes how, in recent years, with the development of microelectronics, implantable electronic devices have been playing a significant role in modem medicine. Examples of such electronic implant devices are, for instance, retinal prosthesis and brain implants. It brings great challenges in low power radio frequency (RF) and analog designs. This article presents a low power Gaussian frequency shift keying (GFSK) demodulator designed for Medical Implant Communications Service (MICS) band Receiver. This demodulator utilizes a novel structure that a wide IF range can be handled and presents the smallest Δf/f ratio in any published GFSK demodulators. In theory the demodulation method can be applied to any RF frequency. The demodulator draws 550uA from a 1 V power supply. A maximum data rate of 400 Kbits/s can be achieved within the 300 KHz channel bandwidth defined by MICS. A simulated signal-to-noise ratio (SNR) of 15.2dB at AWGN channel is obtained to achieve 10^{-3} bit error rate (BER). This demodulator is fabricated on 65-nm CMOS and occupies 0.12mm2 silicon area.

DOI: 10.4018/978-1-7998-1204-3.ch045

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

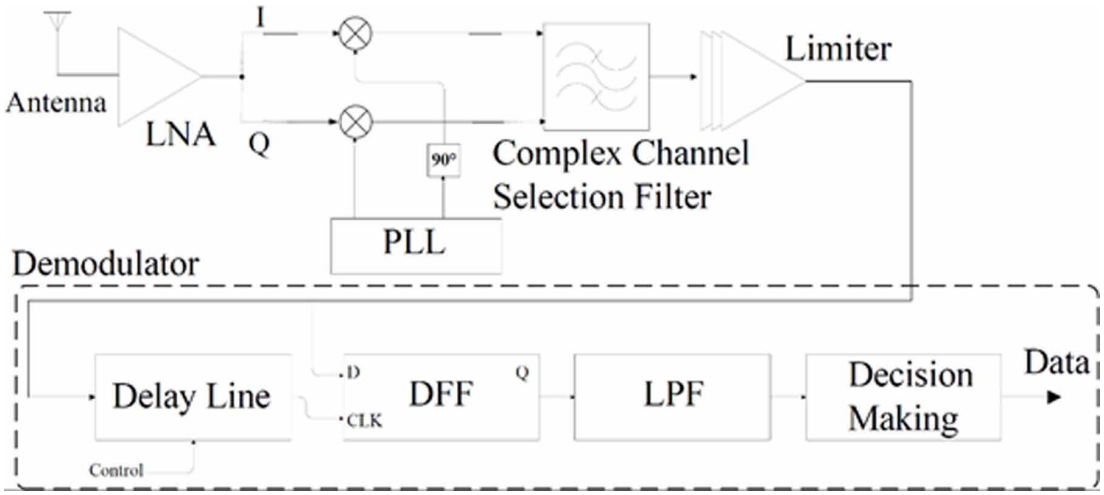
INTRODUCTION

The MICS band is an ultra-low power, unlicensed mobile radio service for implanted medical device. Ten 300 KHz channels are allocated at 402-405 MHz frequency band ("FCC Rules and Regulations," 2002). Implanted wireless device has a requirement of high reliability and ultra-low power consumption. Recent CMOS technology has enabled the design of reliable, low power and cost effective wireless transceiver. This demodulator is designed on 65nm technology under 1 V supply. It intends to be used for the epi-retinal prosthesis (Bionic Eye) to demodulate data for a 32x32 electrode array. In order to support these 1024 electrodes, a data rate more than 300 Kbps is required to make fully functional. Achieving high date rate and low power consumption from this limited 300 KHz bandwidth are challenges in the design of such system. So far, amplitude shift keying (ASK) and binary shift keying (BFSK) have been widely used in biomedical implants (Ghovanloo, Beach, Wise, & Najafi, 2002; Ghovanloo & Najafi, 2003). However, these two methods are not bandwidth efficient. This paper presents a demodulator that can work for both continuous phase BFSK and GFSK modulation schemes. Using GFSK, a data rate up to 400 Kbit/s is achieved within the given 300 KHz bandwidth. This demodulator is also immune to a frequency offset up to 4 MHz. A BER of 10^{-3} is obtained at SNR of 15.2 dB at the input.

DEMODULATOR ARCHITECTURE

The Architecture of the RF receiver is shown in Figure 1. The RF frequency is down converted from MICS band, 402-405MHz to around 5 MHz. A channel-select filter is used to reject the image and select the desired channel. Then the limiter amplifies the IF signal to rail-to-rail swing. Block diagram of proposed GFSK demodulator is shown in the dashed box of the figure. It consists of four blocks: delay line, delay flip-flop (DFF), gm-c low pass filter, and decision-making circuit. The delay element delays the IF input by approximate half the bit period. Then the DFF uses the delayed signal as the clock to sample the original signal. So the frequency information of GFSK signal is mapped into the density and width

Figure 1. Block diagram of receiver with the proposed GFSK demodulator



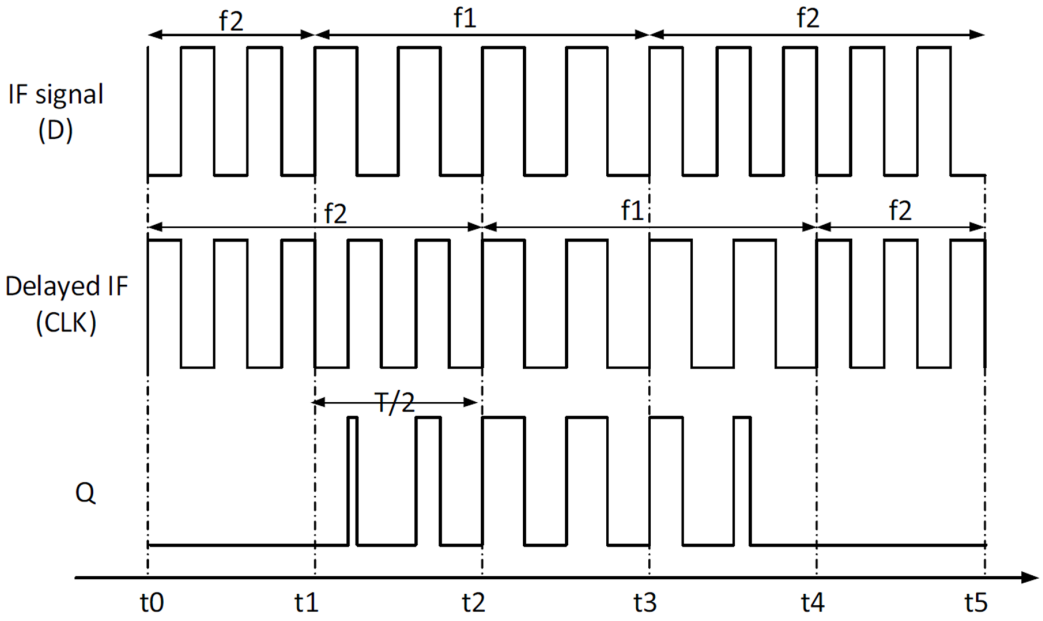
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of the pulse train. After that, the low-pass filter converts the frequency information into voltage level difference. The voltage level after low-pass filter (LPF) may experience variation, as it can be affected by the frequency offset, pulse width variation, and process variation. The decision-making circuit is at the final stage to make a decision on the specific input.

Compared with conventional phase-quantization-based GFSK demodulator (Han & Zheng, 2009) and delay-line based BFSK demodulator, the proposed demodulator can tolerate much larger frequency offset. In other words, it is capable of working at a wider frequency range. Figure 2 shows the principle how this demodulator works at different frequencies. A BFSK signal with bit sequence starting with "010" is used as IF input. Two frequency tones f1 and f2 are used to present logic "1" and "0" respectively. Frequency deviation is Δf, where Δf = f2 - f1. From time t0 to t1, IF signal is sampled by its own delayed signal with same frequency f2, therefore the DFF produces a pulse train with equal pulse width at Q output. Then between time t1 and t2, frequency of original signal, f1, is smaller than that of the delayed signal, f2, so the pulse width at Q output of DFF is drifting. Since t2, both original signal and delayed signal are represented by frequency f1, the output of Q maintain the same width until t2 when the bit transition of original signal occurs. From t3 to t4, the two signals again are having different frequencies. But this time IF signal is sampled at a lower frequency f1. This is a reversed process of the pulse width drift happened between t0 and t1.It can be concluded that output pulse width is fixed at the center of each bit period, pulse width drift occurs around bit transition. Thus, frequency information of FSK signal is contained in the pulse train. When GFSK signal is applied, a similar output would be resulted except that the width of pulse changes continuously. Simulation result with GFSK input will be shown in the last section.

An important feature of this technique is that it is suitable not only for IF but also RF band, while other published works such as Han and Zheng (2009) and Xia et al. (2003) only work for low IF application as they require a relatively large ratio of bandwidth over carrier frequency to work properly. Same as in

Figure 2. Operation concept of self-sampling



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fu et al. (2010), this demodulator uses a low-pass filter to convert frequency information into voltage variation. However, in this work, the big bandwidth-to-carrier frequency ratio is not necessary. After LPF power averaging, DC voltage level variation between bit "1" and bit "0" could be expressed as:

$$V_{var} = \frac{n\left(\frac{1}{2f_1} - \frac{1}{2f_2}\right)}{\frac{1}{f_1}}, \quad \mathrm{n} = \frac{\frac{T}{2}}{\frac{1}{f_1}}.$$

where *n* is the number of pulses during the half bit period of the original signal. Also,

$$\frac{1}{f_1} - \frac{1}{f_2} = \frac{f_2 - f_1}{f_1 f_2} = \frac{\Delta f}{f_1 f_2} \approx \frac{\Delta f}{f_1 f_1}.$$

So, voltage variation can be re-written as:

$$V_{var} \approx \frac{n\left(\frac{\Delta f}{2f_1 f_1}\right)}{\frac{1}{f_1}} = \frac{\frac{\frac{T}{2}}{\frac{1}{f_1}}\left(\frac{\Delta f}{2f_1 f_1}\right)}{\frac{1}{f_1}} = \frac{\Delta f T}{4}$$

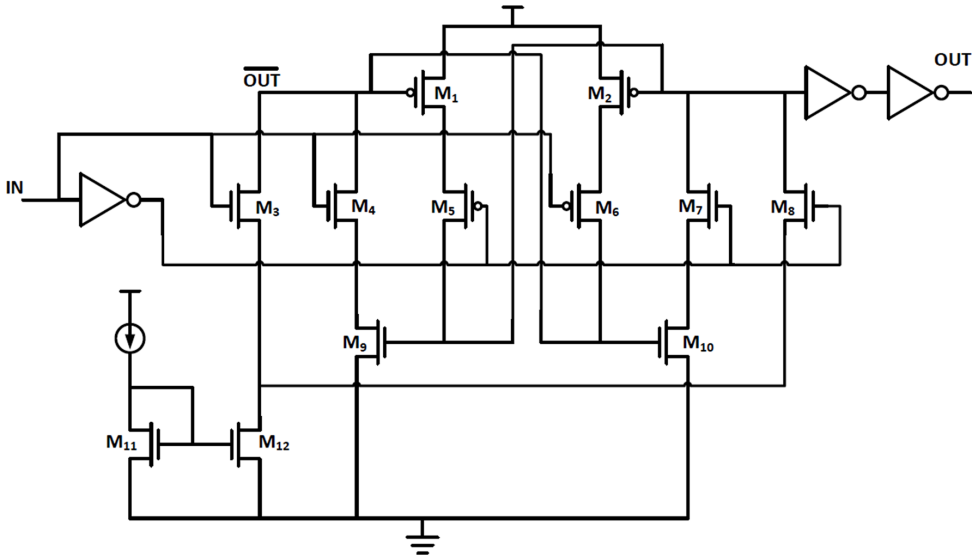
So, it can be concluded that the DC voltage variation is only dependent on Δf and bit period T. The only limit of using this demodulator for RF band is due to the difficulty providing large delay at a higher frequency.

DEALY ELEMENT

A delay element is a circuit that introduces a delay to the input signal and produces a lagged waveform with a similar shape at the output. The delay element used in this work is based on the concept of thyristor (Kim, Kim, Chang, & Kim, 1996). Schematic of this thyristor delay element is shown in Figure 3. Transistors *M1, M3, M4, M5* and *M9* are involved with delaying the rising edge of the input, while falling edge is delayed by *M2*, *M6, M7, M8* and *M10*. Operation concepts of rising edge delay and falling edge delay are exactly the same. Falling edge delay is indeed accomplished by delaying the rising edge of the inverted input signal. So only the rising edge delay operation needs to be discussed. When input changes from low to high, it switches on *M3* and *M4* and starts discharging 'OUT' through the pull-down network composed of *M11* and *M12*. Then, after 'OUT' drops below (V_{DD} - V_{Tp}), *M1* is switched on. *M5* has already be turned on since 'IN' went high. Once both *M1* and *M5* are on, they start charging up 'OUT'. Then 'OUT' reaches to *M9*'s threshold voltage V_{Tn}, turns on *M9* and pulls further down the voltage of 'OUT' until it drops to Vss. During this process, *M2, M6, M7, M8* and *M10* all stay in 'OFF' state. The

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The CMOS thyristor as delay element



delay time is determined by the current source, thus can be easily controlled. Thyristor based delay element provides good signal integrity (Kim et al., 1996). Because of the positive feedback configuration, relatively sharp rising and falling edges can be obtained. Note that overall delay time of any delay cell cannot be larger than the pulse width. Multiple delay elements are required to provide enough delay.

TRANSCONDUCTANCE-C LOW PASS FILTER

The low-pass filter is a key block in this demodulator. It averages the power in the pulse and converts the pulse density that carries the frequency modulation information, into a voltage variation. Passive devices are traditionally adopted for the low-pass filter at high frequency. However, at a low cut-off frequency, conventional LC low-pass filter occupies too much area and it is not suitable to build on the chip. Nowadays, active transconductance-capacitor (also referred to as gm-c) filters have become more prevalent and are widely used for on-chip integration.

Various transconductance (OTAs) have been applied in gm-C filter design. Natua's OTA is chosen in this application, because of its good linearity and infinite bandwidth (1992). The structure of Nauta's OTA is shown in Figure 4. Assume that nfet and pfet of an inverter are perfectly matched by choosing appropriate size of the transistors. pfet is about three times larger than nfet to compensate the lower electron mobility. When both pfet and nfet are operating in saturation region, the drain current can be analyzed as follow:

$$I_{dn} = \tfrac{1}{2}\,\beta_n\,(V_{gsn}-V_{tn})^2, \text{ with } \beta_n = \mu_n\,C_{ox}\,W_n/L_n$$

$$I_{dp} = \tfrac{1}{2}\,\beta_p\,(V_{gsp}-V_{tp})^2, \text{ with } \beta_p = \mu_p\,C_{ox}\,W_p/L_n$$

Subtraction of I_{dn} and I_{dp} would give the output current:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$I_{out} = I_{dn} - I_{dp} = a\,(V_{in} - V_{tn})^2 = b\,V_i\,n + c$$

where

$$a = \tfrac{1}{2}\,(\beta_n - \beta_p)$$

$$b = \beta_p\,(V_{dd} - V_{tn} + V_{tp})$$

$$c = \tfrac{1}{2}\,\beta_p\,((V_{tn})^2 - (V_{dd}+V_{tp})^2)$$

For perfect matched pfet and nfet, $\beta_n = \beta_p$ and current of the n-channel device equals current of the p-channel device. When both devices are in saturation, the output current is 0. Threshold voltage of the inverter, also known as common mode voltage, is equal to 1/2 V_{dd}. The first level of inverters perform the basic V-to-I conversion. Without the second level inverters, the output current of inverter 1 and inverter 2 can be calculated and the value of differential output current I_{od}:

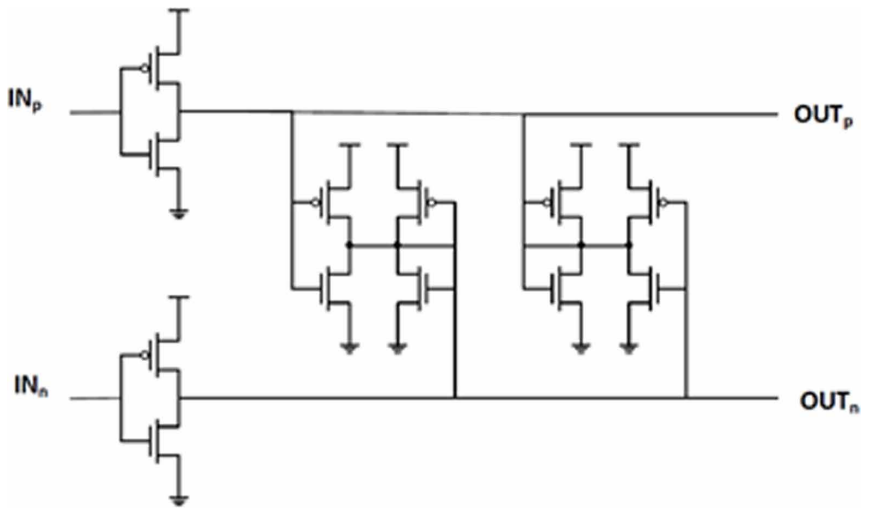
$$I_{o1} = a\,(V_c - V_{tn} + \tfrac{1}{2}\,V_{id})^2 + b\,(V_c + \tfrac{1}{2}\,V_{id}) + c$$

$$I_{o1} = a\,(V_c - V_{tp} + \tfrac{1}{2}\,V_{id})^2 + b\,(V_c + \tfrac{1}{2}\,V_{id}) + c$$

$$I_{od} = I_{o1} - I_{o2} = V_{id}\,(\beta_p\,(V_{dd} - V_c + V_{tp}) + \beta_n\,(V_c - V_{tn}))$$

So, the differential output current I_{od} is linear with differential input voltage, $V_{id}=V_{inp}-V_{inn}$. It can be approved that the linear relationship is valid even with imperfect matched pfet and nfet. Another import feature of Nauta's OTA is that it achieves infinite bandwidth voltage to current conversion. Note that the first level of inverters performs voltage to current conversion with no internal nodes. The second level inverters are used to control DC gain and do not present any other internal nodes as well (Nauta, 1992).

Figure 4. Nauta's transconductance



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Fifth order Butterworth bandpass low-pass filter prototype

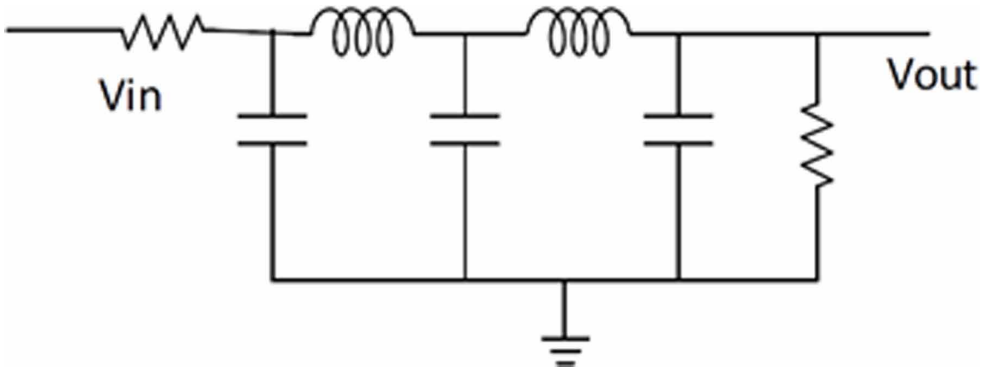
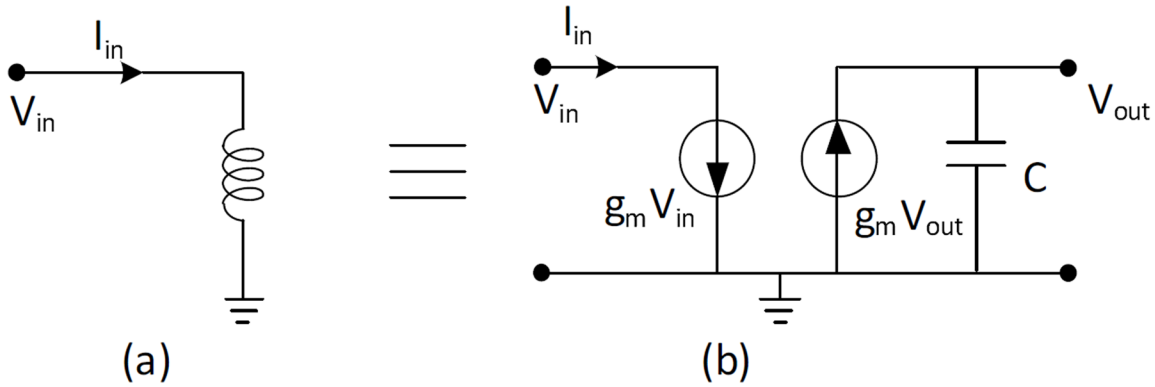


Figure 6. a) A grounded inductor and b) realization of the grounded inductor with gyrator



As shown in Figure 5, a fifth order bandpass low-pass Butterworth filter is selected as the filter prototype for its small ripple. The cut-off frequency is designed to be 200 KHz. By using replacement method of LC ladder, each inductor in the filter prototype can be replaced with a gyrator-capacitor circuit. The transfer function of an inductor is equivalent to capacitively loaded gyrator. Figure 6 shows the principle of LC ladder simulation of a grounded inductor. Similarly, a floating inductor can be simulated with two gyrators loaded by a capacitor in the middle. When the relationship between inductor and gyrator-capacitor can be written as

$$L = C/g_m^2$$

DECISION MAKING CIRCUIT

Then DC voltage level after low-pass filter could be affected by frequency offset, delay and process variation. In order to eliminate the effect of varied DC voltage level, a final decision should be made on the specific input. Figure 7 presents the architecture of proposed decision-making circuit (Koli & Halonen, 1997). A peak detector and a valley detector (Han & Zheng, 2009) detect and save the maximum

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

voltage, V_{peak}, and minimum voltage, V_{valley}, of the incoming signal separately, and then stores them in capacitors. A voltage divider is used to find the average voltage of V_{peak} and V_{valley}, where

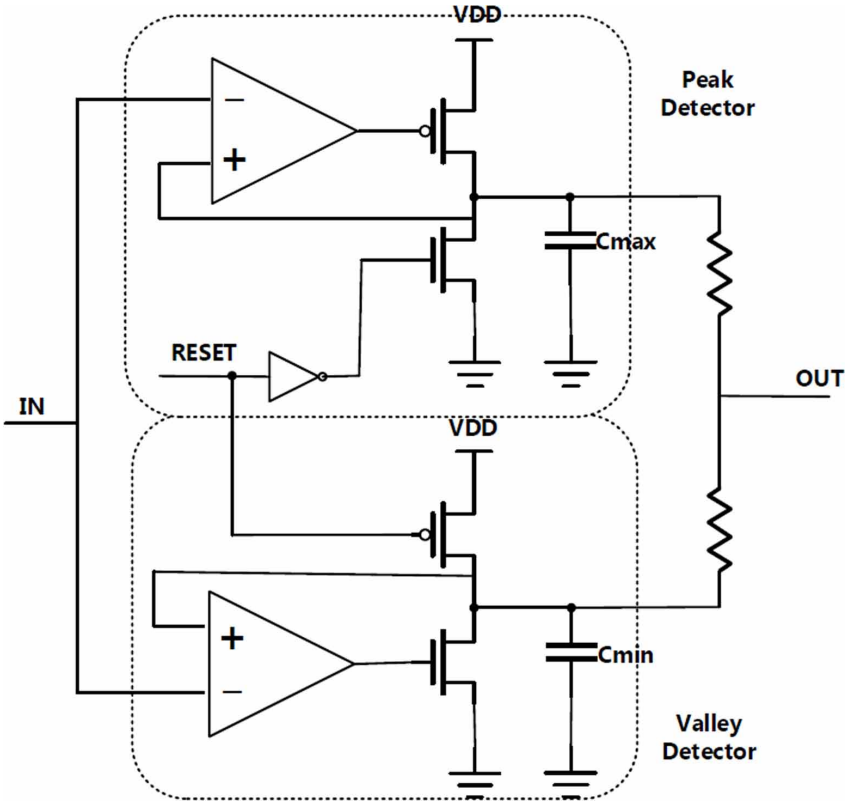
$$V_{avr} = (V_{peak} + V_{valley}) / 2$$

A final decision could be made by comparing the input with V_{avr}. A "0" is obtained when the input is below V_{avr} and a "1" is obtained when the input is larger than V_{avr}.

BIT SYNCHRONIZER

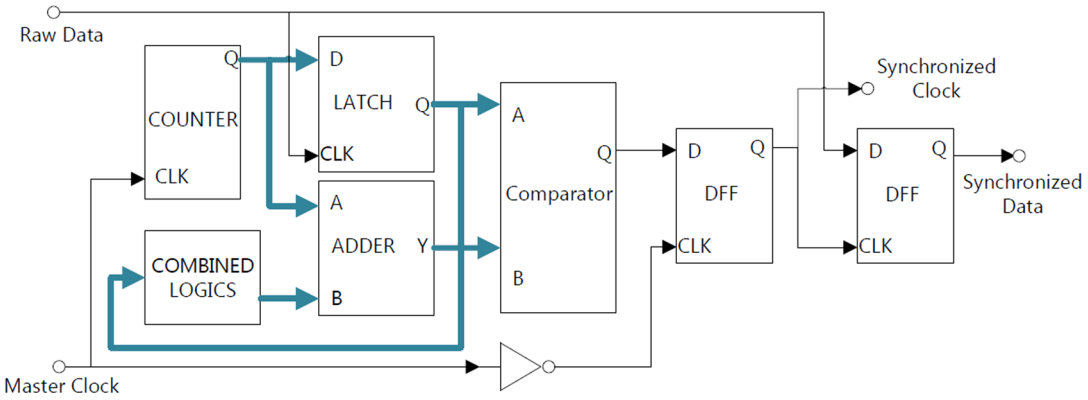
After decision making, demodulated data will not have 50 percent duty cycle. When simulating the demodulators with process variation (Monte Carlo simulation) and at different corners (such as FF and SS corners), the duty cycle can vary between 30 percent and 70 percent. Also, the bit stream is not synchronized with the system clock. This bit stream with un-perfect duty cycle is fed to the bit synchronizer, which will synchronize the bit stream to the system clock and reshape it into 50 percent duty cycle. The bit synchronizer is first reported in Piedade et al. (2005). As shown in Figure 8, it consists of a counter, a 4-bit latch, a 4-bit adder, two D flip-flops, a comparator and some logics.

Figure 7. Peak detector, valley detector and decision making



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Peak detector, valley detector and decision making



The master clock frequency is 4MHz, which is ten times of the bit rate of output from the demodulator. Rising edge of the first input bit will trigger the counter. Then it starts counting the master clock. The counter is able to count from0 to 9. When it counts to 4, the raw date is being sampled. There are ten master clock cycles within each bit period. So sampling point is in the middle of the first bit. After first bit is sampled, the counter will be reset and continue counting another 10 cycle regardless whether the raw data is 0 or 1. After bit synchronizer, demodulated data is re-shaped with 50% duty cycle and synchronized with the master clock.

SIMULATION RESULTS

This section presents the simulated performance of proposed GFSK demodulator. GFSK signal with *100KHz* frequency deviation and additive white Gaussian noise is applied to the demodulator. Figure 10 shows the signal with *14dB* SNR. Peak to peak voltage swing at input of the demodulator is *300mV*. 10^{-3} bit error rate is achieved at signal noise ratio of *15.2dB*. A *6μS* delay is introduced during the process of demodulation. The group delay remains constant over bit transition. Figure 9 shows the layout of proposed demodulator. Simulation results are present in Figure 11, 12, 13, 14. The demodulator consumes *550μA* from a *1V* power supply and occupies 0.1mm^2 silicon area.

CONCLUSION

A low power, wide input frequency range GFSK demodulator in 65nm CMOS was designed and simulated with post-layout extraction. The self-sampling technique greatly improves the tolerance of frequency offset and relaxes the sensitivity requirement of final decision making. With low power consumption, sufficiently wide input range, and MICS band operation, this demodulator is well suited for the retinal prosthesis and other biomedical implanted devices. The novel demodulation method also makes PLL-less receiver possible for low-power biomedical applications. This demodulator is part of a sub-1mW PLL-less MICS band receiver.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Layout of proposed GFSK demodulator

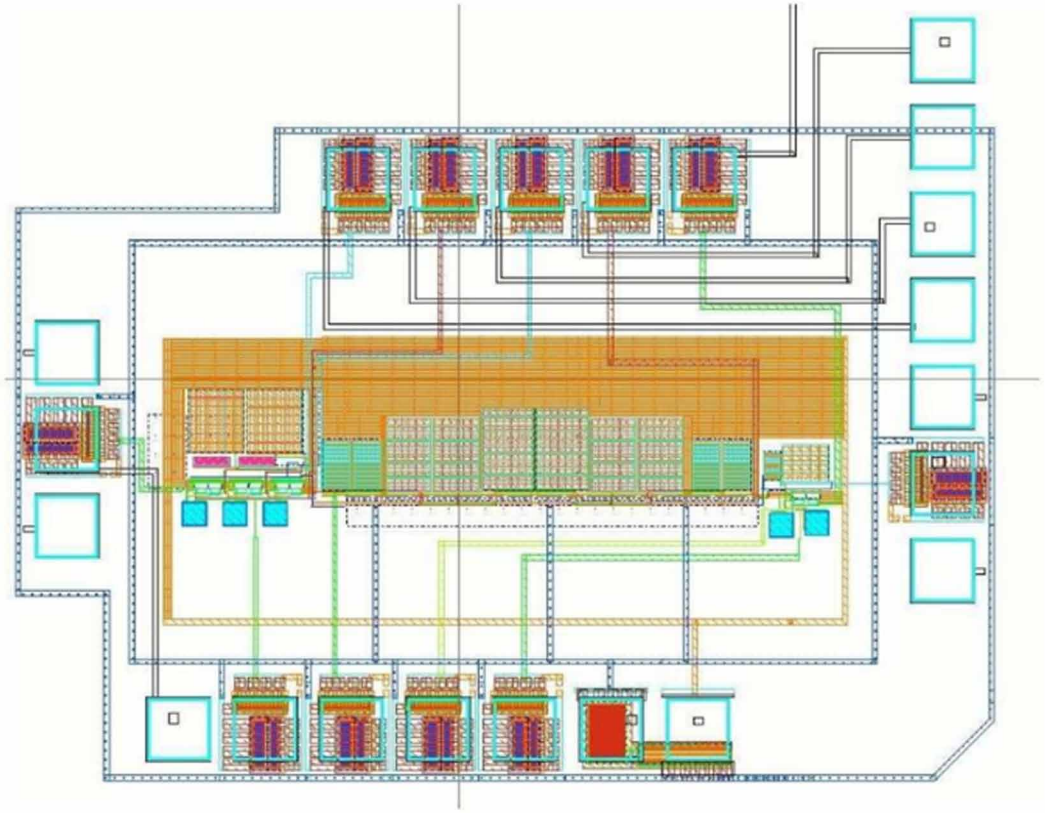
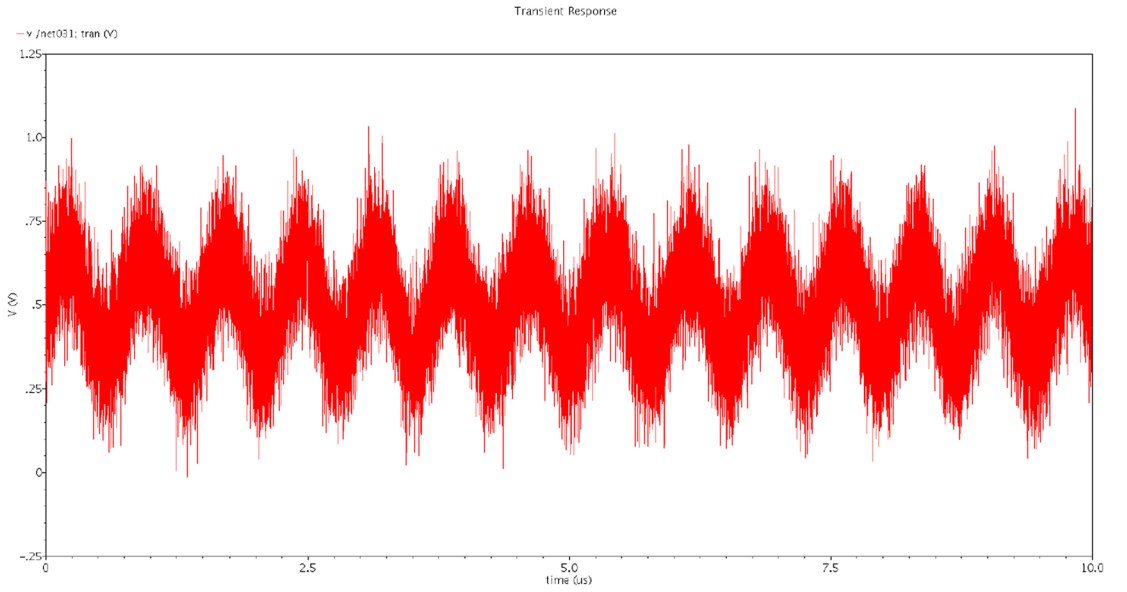


Figure 10. Signal with additive white Gaussian noise



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Modulating signal

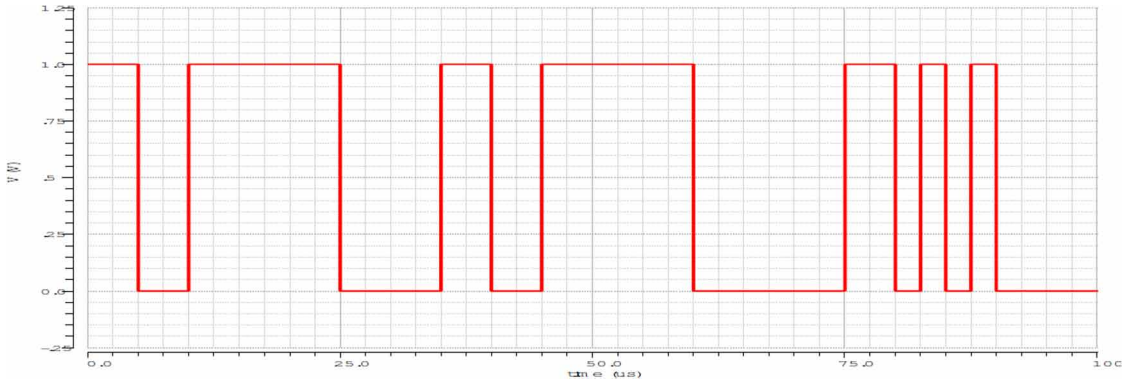


Figure 12. Modulating signal after Gaussian shaping

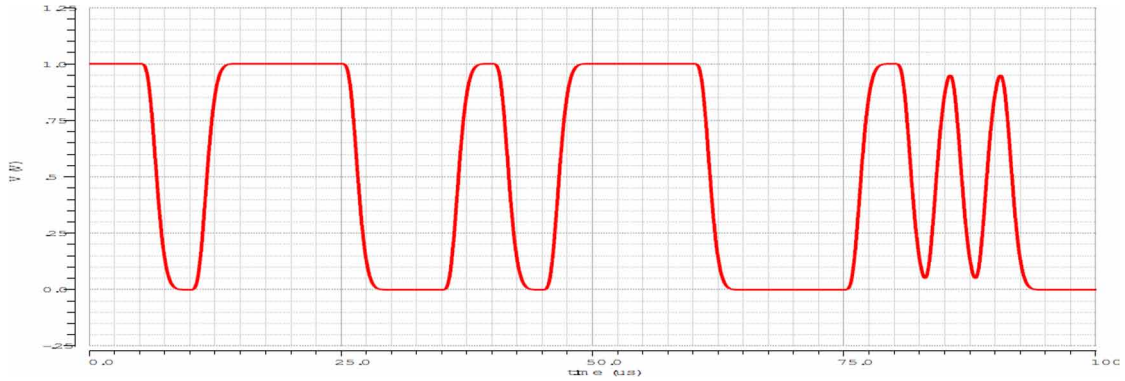
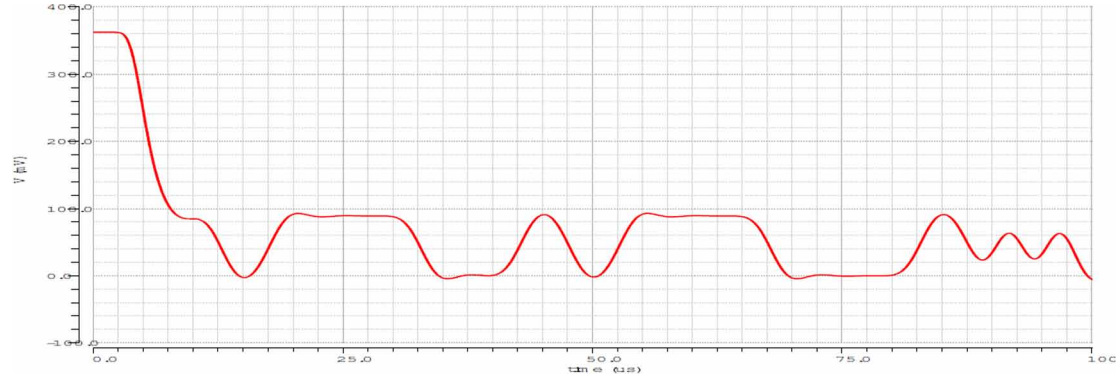
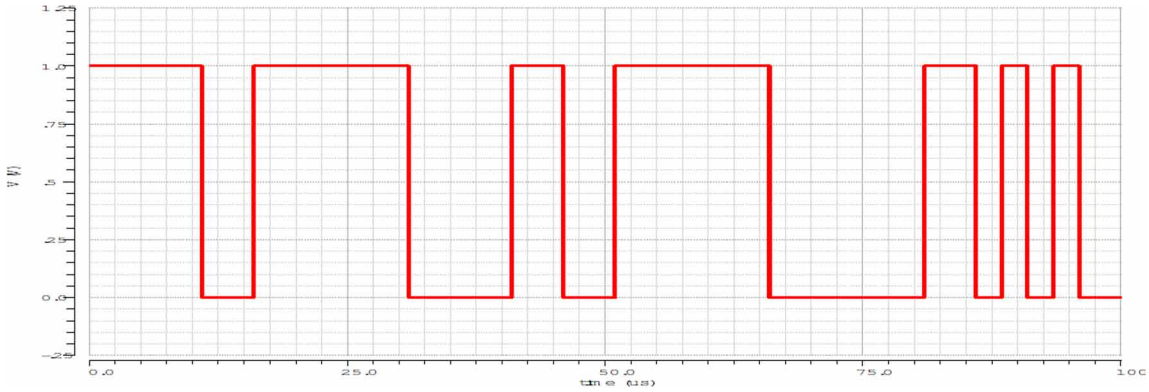


Figure 13. Peak detector, valley detector and decision making



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Demodulated data after decision making



REFERENCES

Fu, M., Yang, J., Tran, N., Bai, S., Ng, D. C., Halpern, M., ... Mareels, I. (2010). A GFSK demodulator for ultra-low power MICS bandreceiver. In *Proceedings of the 5th International Conference on Broadband and Biomedical Communications*, Malaga.

Ghovanloo, M., Beach, K., Wise, K. D., & Najafi, K. (2002). A BiCMOS wireless interface chip for micromachined stimulating microprobes. In *IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology proceedings*, Madison, WI (pp. 277-282) . 10.1109/MMB.2002.1002330

Ghovanloo, M., & Najafi, K. (2003). A fully digital frequency shift keying demodulator chip for wireless biomedical implants. In *Proceedings of the Southwest Symposium on Mixed-Signal Design* (pp. 223 – 227). 10.1109/SSMSD.2003.1190431

Han, D. & Zheng, Y. (2009). A Mixed-Signal GFSK Demodulator Based on Multithreshold Linear Phase Quantization. *IEEE Transactions on Circuits and Systems*, *56*(9).

Kim, G., Kim, M.-K., Chang, B.-S., & Kim, W. (1996, July). A Low-Voltage, Low-Power CMOS Delay Element. *IEEE Transactions on Circuits and Systems*, *31*, 966971.

Koli, K., & Halonen, K. (1997). Low voltage MOS-transistor-only precision current peak detector with signal independent discharge time constant. In *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS'97)*, Hong Kong (Vol. 3, pp. 1992-1995). 10.1109/ISCAS.1997.621544

Nauta, B. (1992). A CMOS transconductance-C filter for very high frequencies. *IEEE Journal of Solid-State Circuits*, *27*(2), 142–153. doi:10.1109/4.127337

Piedade, M., Gerald, J., Sousa, L., Tavares, G., & Toms, P. (2005). Visual neuroprosthesis: A noninvasive system for stimulating the cortex. *IEEE Transactions on Circuits and Systems. I, Regular Papers*, *52*(December), 2648–2662. doi:10.1109/TCSI.2005.857923

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

FCC Rules and Regulations, Table of Frequency Allocations, Part2.106. (2002).

Xia, B., Xin, C. Y., Sheng, W. J., Valero-Lopez, A. Y., & Sanchez-Sinencio, E. (2003, August). A GFSK demodulator for low-IF Bluetooth receiver. *IEEE Journal of Solid-State Circuits*, *38*(8), 1397–1400. doi:10.1109/JSSC.2003.814424

This research was previously published in the International Journal of Interdisciplinary Telecommunications and Networking (IJITN), 10(3); edited by Steven R. Powell and Michael R. Bartolacci; pages 21-32; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 46
Step Towards Interoperability in Nursing Practice

Daniela Oliveira
University of Minho, Portugal

Júlio Duarte
University of Minho, Portugal

António Abelha
University of Minho, Portugal

José Machado
University of Minho, Portugal

ABSTRACT

Hospital inpatient care compromises one of the most demanding services in health institutions for providing a careful and continuous healthcare assistance. Such demands require a constant update of the patients' health record allied with support systems responsible for monitoring their clinical information. In this context, the problem in this study becomes a process of continuous improvement. To define the case study, it was necessary to use research tools such as questionnaires and interviews. With these techniques, it was possible to delineate the state and dimension of the problem. Subsequently, the approach and solution was established and a new web platform for the daily monitoring of patients was proposed focused on nurses. The tool incorporates a real-time data visualization, and a patient record during an inpatient care episode. Moreover, this article also highlights the required adaptability of this platform for each health unit according to needs. With this solution, it is expected to correct many of the problems detected through quantitative results.

DOI: 10.4018/978-1-7998-1204-3.ch046

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

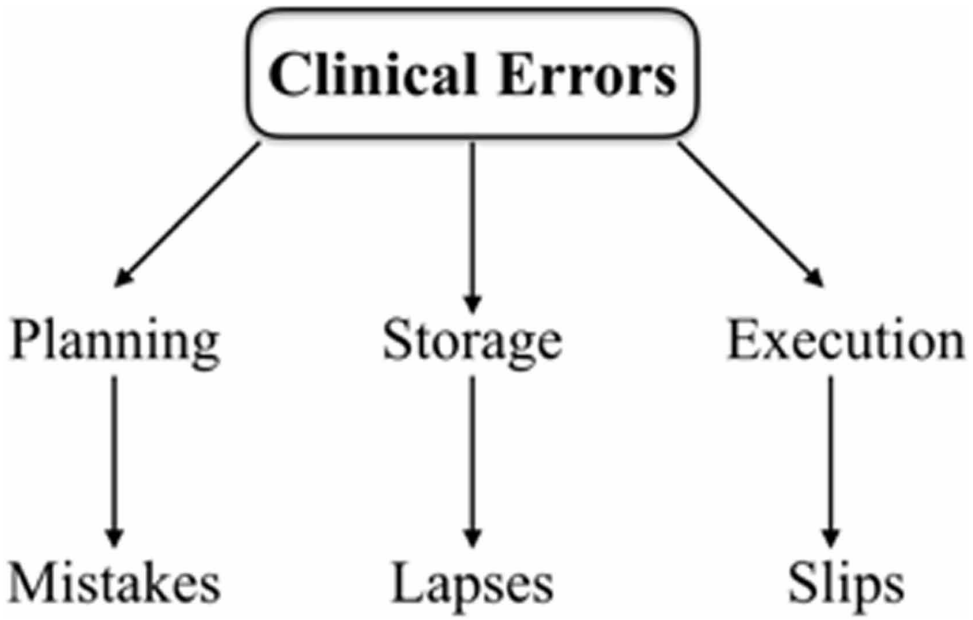
INTRODUCTION

In a hospital environment, the patient security is one of the major concerns in health, in the sense of improving the efficiency of healthcare delivery, and minimizing delays or failures that could lead to other consequences. These consequences can be translated to a considerable portion of clinical errors which can be of several types, depending on the approach taken. Regarding effects, Reason (2000) argues that errors can be active, with an immediate or latent impact. Thus, the results can appear in short-term or long-term, respectively. According to Reason, these errors can have different sources, disclosed in Figure 1. Moreover, James Reason's "Swiss cheese" model of error causation defines that in complex organizations such as hospitals, errors occur due to the holes or flaws in the multiple layers of defense, causing adverse outcomes to the patients (Reason 2010, Salazar et al., 2013, Duarte et al., 2011, Portela et al., 2014).

The Institute of Medicine's Quality of Health Care in America valued that about 98000 people every year die due to clinical errors at hospitals. Comparing to other main causes of death in America, this causes a greater impact in the scientific community, becoming one of the most emerging public problems today (Donaldson 2000).

The continuous and dynamic alteration of patients' clinical status in healthcare institutions is one of the major causes of the occurrence of adverse events. This level is very high becoming urgent to find connections with other areas, such as information technologies (IT) (Peixoto et al 2012). This is one of the most explored fields in medical informatics, raising great challenges in terms of development, implementation and maintenance. Thus, over the years, and due to the high data flow, efficient and user-friendly support systems for health professionals have been emerging, called hospital information systems (HIS) (Coiera 2003).

Figure 1. Classification of clinical errors according to (Reason, 2010)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The high number of professionals involved in clinical acts, combined with their diversity of academic backgrounds, has led to high level of automation of services, from the well-known clinical software, which has several advantages such as financial, administrative, and recruitment, until reaching the medicine itself. There, the role of these developed HIS is crucial (Khodambashi 2013, Machado 2007). Its main goal is to streamline and support the provision of direct healthcare to the patient, being some of them oriented to the use of a certain group of professionals, such as doctors and nurses (Kim & Park 2011).

In Europe, the Health2020 strategic program organized by World Health Organization (WHO) highlights nurses and their important role in improving healthcare delivery, since they are the most direct and continuous contact with the patient throughout the hospitalization episode. The essence of nursing is "caring", i.e., providing the highest possible quality care for an assertive practice. As such, these professionals are responsible for providing safe, high quality and efficient healthcare services. To provide this high quality, all the conditions must be attended. Thus, one of the main focuses of this program is the nursing team (World Health Organization 2015, Duarte 2011, Rocha & Rocha 2014).

To overcome failures in shifts change, such as administration of a given medication with the wrong dosage or omission of some relevant information about the patient, it emerged the motivation to develop a nursing care support system in the hospitalization. This platform aims to automate information for an easy visualization, excluding the need to resort to other means that would certainly take more time. All the patient's relevant information for nurses can be consulted on this platform, such: the location of his/her room; identification of number of the bed; the type of fasting that is attributed to him with a specific examination or analysis that he/she will carry out; the alerts and medication associated with the patient; among other attributes. This data is updated continuously.

The main contributions with this work are:

- Potential reduction of occurrence of clinical errors (medication errors, dosage errors, diagnostic errors, among others) committed through clinical information made available in real time;
- Support health professionals during the hospitalization episode, gathering relevant and updated information about the patient, introducing relevant attributes before the clinical act;
- Increase the efficiency of health care provided;
- Improve the quality of patient care;
- Application of the platform developed in the various hospital services, being initially tested in the Hospital Center of Porto (CHP) Cardiology service.

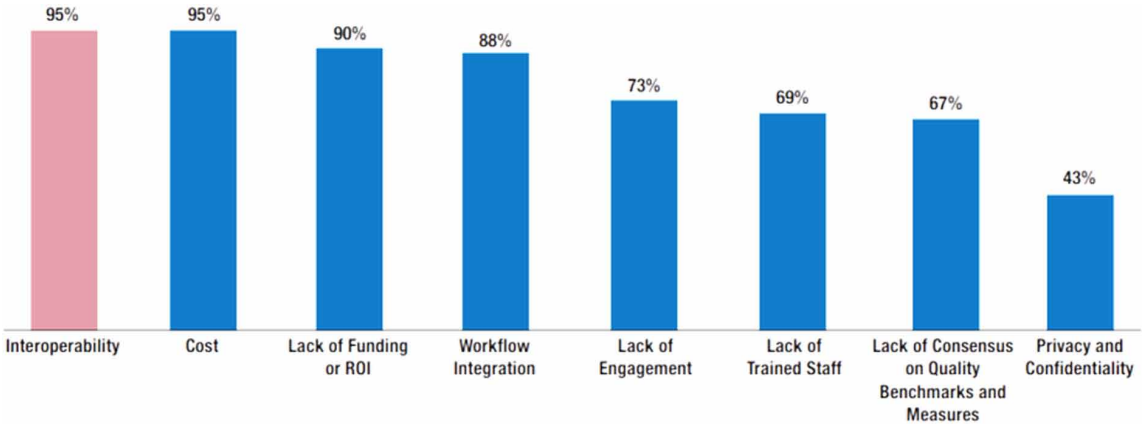
The presented paper is organized as follows: the next section describes the importance of the interoperability of systems and how this can be guaranteed through multi-agent technology. Thereafter, depicts the research methodologies used in the project as well as main features and advantages of the web platform developed and discusses its relevance respectively. Finally, a main conclusion is outlined according to this work.

INTEROPERABILITY AND MULTI-AGENT SYSTEMS

The presentation of clinical information is done in several ways, depending on the kind of data stored in the systems. For this reason, it often becomes difficult for health professionals to integrate quickly into the patient's condition.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Percent of ACOs reporting largest barriers to using health IT effectively, 2014 (American Hospital Association, 2015)



In this context, a USA study shows the current state of interoperability in the country. Of the 62 Accountable Care Organizations (ACOs), health organizations according to the American health model, 95% referred to interoperability as one of the main obstacles to harnessing the true potential of IT in health (American Hospital Association 2015) (see Figure 2).

For all HIS to operate at their maximum efficiency and to relate to each other without communication failures and with high coherence, the concept of interoperability takes great importance in this context (Cardoso 2014).

In the middle of 1960s, the Healthcare Information and Management Systems Society (HIMSS) emerged a non-profit association, whose mission is to unite efforts to improve health through information technology continuously (Duarte 2016). Following their definition "Interoperability describes the extent to which systems and devices can exchange data, and interpret that shared data. For two systems to be interoperable, they must be able to exchange data and subsequently present that data such that it can be understood by a user…" (HIMSS, 2017). Nowadays, a high percentage of relationships between HIS still lacks interoperability among them. This is caused by the high development and complexity of the current databases that support hospital systems. There are several advantages provided by interoperability in health, among them the reduction of the probability of occurrence of clinical errors and consequent improvement of the quality of the services provided (Dove, 2005). The exchange of useful information between different health institutions has proved to be a necessity. For example, after a period of inpatient care, the patient is followed by a general practitioner and the electronic clinical record must be properly updated (Lang, 2006). Clinical decision support systems also contribute to improved health care delivery by helping health professionals to make their daily decisions (Berner & Moss, 2005, Moreira et al., 2017). These realities, among others, reveal the need for constantly updated and easily accessible clinical information. These features can be ensured through investment at the various levels of interoperability. One of the most common and effective solution used to solve the miss of interoperability are multi-agent systems (MASs). These systems consist of an environment created by the interoperation of two or more agents that make adaptable but standardized information flow enabling interoperability in a loosely coupled manner (Miranda *et al*. 2012). Moreover, they are associated with data warehouses for efficient data storage and query. For that matter, this was adopted in this project.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Multi-Agent Systems

Currently, the volume of clinical data is very high and there has been a need to build consistent and organized data models, such as Data Warehouses (DWs). DWs are subject-oriented, integrated, time-variant, and non-volatile collections of data in support of management's decision-making process (Inmon, 1996). To coordinate and update their content, multi-agent technology was used.

Being multi-agent architectures a field of research of distributed artificial intelligence, this technology is intrinsically connected to the concepts that define a distributed architecture, while being distinct in the definition of an agent versus the properties of the general middle-wares of many others distributed architectures. Being distributed by nature its characteristics define MAS as an abundant and fiercely adaptable technology with great interest mainly due to the research interests in this area (Miranda et al., 2012). A MAS consists of a computer system with multiple autonomous agents that cooperate together to perform certain tasks or achieve certain objectives. An agent can be classified as homogeneous or heterogeneous and is defined as an element able to respond autonomously to problems, operating asynchronously, always respecting other existing agents. Its main characteristics are the autonomous capacity to make decisions, as well as the capacity to interact with other agents through social interaction protocols, reaching the desired levels of coordination, cooperation, competition and negotiation.

Study Considerations

The main goal of this study is to automate the hospital information, updating it constantly, without any effort or action that might interfere with the process of healthcare delivery. To accomplish this purpose, a web application was developed, dedicated to nurses, able to show the inpatient's information in health units. Since we are dealing with big data, the solution adopted addressed the use of multi-agent systems for monitor and manipulate the developed DW. This DW is suitable for a system that is capable of crossing several kinds of information from different sources (Machado et al., 2007, Machado et al., 2006). With the DW built, bridges between different HIS, previously non-existent, are established. Regarding the current study, some information was transformed and adapted to the project, given the needs of the users of the platform. The operation of the developed MAS will be described posteriorly, as well as the strategy established so that the DW information is constantly assured and updated. At this stage of the project, the interoperability between the systems used is guaranteed, reducing some possible communication failures at any level (Duarte et al., 2010).

METHODOLOGIES OF RESEARCH

In the initial phase of the project, several studies and tests were carried out in order to understand the needs of the project. Afterwards, the results obtained were analyzed for a later validation of the best model and architecture. Two methodologies and three research techniques were used, as will be presented.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Case Study

A case study (CS) is a methodology which contributes to the definition of research topics, for a given context. Among several methodologies used in information systems, the CS was chosen for the definition of the project, given its highlight among other strategies, especially in the organizational context.

Pozzebon & Freitas (1998) claimed that the researcher must define clearly the initial goals and the reason for the choice of the research target, and describe in detail the techniques and sources used for the analysis and collection of data, avoiding some common mistakes detected in the application methodology. Additionally, the researcher must be impartial, ignoring his/her personal opinions that can interfere with the results. Moreover, according to (Benbasat et al., 1987), CS has as main characteristic the versatility, which can be used in distinct approaches. Thus, they defined the following eleven points, that will be respected in this work:

- The object of study is evaluated in its natural context;
- The collection data is performed in several ways;
- A few entities are analyzed (groups, persons or organizations);
- The complexity of the unit is studied intensively;
- CS is more suitable for exploration, classification and definition of several stages of developing, being a process of knowledge;
- The research stage does not include control or experimental manipulation;
- The independent and dependent variables might not be defined a priori;
- The results are dependent on the type of integration;
- During the gathering of hypotheses, changing types can occur (local or methods of collecting data);
- CS is suitable to respond questions like “how” and “why”;
- The focus lies in real and actual happenings.

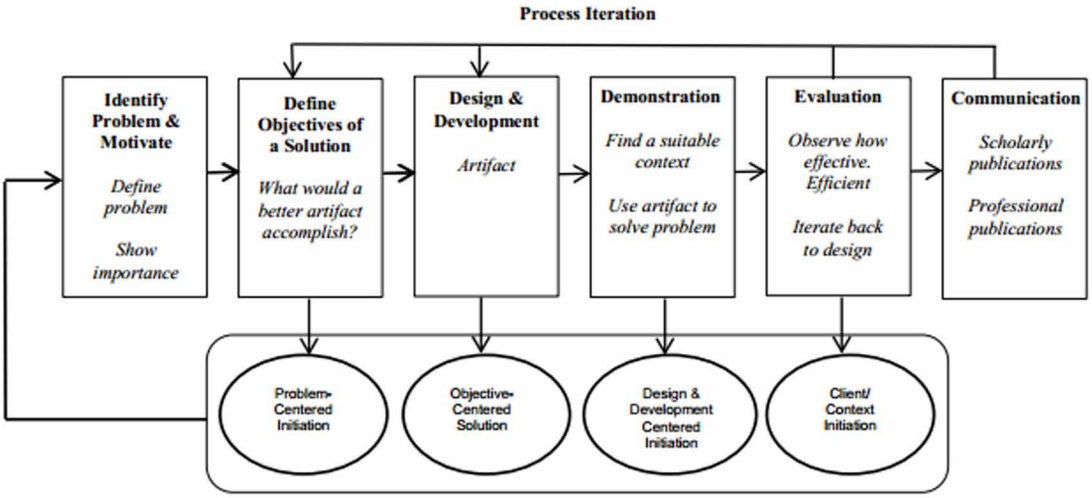
Design Science Research

The design science research (DSR) model used was the one proposed by (Peffers *et al*. 2007) with the structure shown in Figure 3.

The starting point was the identification of the problem. The problem that arises is how the whole relevant clinical information of each patient at the time of hospitalization can be provided at the same place. This information is dispersed by several systems and can be stored in several ways. The goal is to find a solution to the above problem, which helps health professionals to have access to the information they need at a given moment, in a fast and intuitive way. Thus, the chosen solution was to develop an artefact capable of showing all the necessary and constantly updated information about each patient. This artefact in production will be subsequently implemented, firstly in the Cardiology health unit of CHP. Afterwards, the next goal is the its propagation to the remaining units. After its implementation, its impact will be reassessed by the users with specific models.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. DSR methodology process model (Peffers et al., 2007)



Techniques of Research

The methodologies were implemented using research techniques, with the objective of searching and collecting data.

Literature Review

An exhaustive bibliographic analysis was carried out in search of similar existing projects. This stage is very important in all research projects, not only for a better definition of the case study, but also to understand with greater precision the developments of the subject in question, now. According to Cardoso (Cardoso et al., 2010), all investigators must in detail analyze the work of others who proceeded and only after that, proceed to their own adventure. The constant evolution of science must be considered. A good review leads to an easier delimitation of the problem under study and to the discovery of new lines of investigation, avoiding possible repetitions and proposing future goals.

No support platform for hospital inpatient care was found, with specifications that will be described later. However, there are several developments in this area, such as the hospital management system (Bento et al., 2017), which consists of the classification of patients with different degrees of care (minimum, intermediate, high dependency, semi-intensive and intensive) for the organization of appropriate care and sizing of the required nursing professionals.

Interviews

The interviews are a relevant technique in the research process. The involvement of future users provides an accurate feedback about their needs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The researcher can feel the reality and clarify the objectives, the structure that intends to follow during the interview and its goal, earning the trust of the interviewing to assure its collaboration in the future (Lavrakas, 2008).

To accomplish these stage, several visits to the health unit of Cardiology of CHP were done, more precisely, to the nursing room. The chef-nurse was the main person interviewed, who showed great collaboration and interest in the proposed platform. During the interviews, it was reported the major difficulties that nurses from that unit feel daily and discussed the main strategies to the outcome such problems.

Questionnaires

The questionnaires technique is very similar to the interviews with the difference that the interviewer does not has direct contact with the interviewing, and they answer the questionnaire anonymously. Regarding the type of questions, the researcher must be careful since they should be clear, coherent and with maximum neutrality (Williams 2006).

In the first phase of the project, a questionnaire with eight questions was carried out. The group of interviewees where some nurses from the cardiology department, with a final result of 35 questionnaires answered.

Objectivity and simplicity were concerns considered when the questions were constructed, aiming to understand the main failures that respondents have in their workplace.

Among the eight questions addressed, four were highlighted:

Question 1: In your opinion, patients are well monitored throughout the shifts?
Question 2: Are the changes in the clinical status of the patient known to all nurses in the health unit?
Question 3: Have you ever faced with the possibility of complications of the patient's condition due to information system failures?
Question 4: Have you ever been confronted with medication / dosage errors?
Question 5: Have you ever witnessed the occurrence of communication failures, for example, about changes in patient accommodation (bed changes)?
Question 6: Have you ever faced with communication failures, particularly in the absence of records from the previous shift?
Question 7: Have you ever had difficulty accessing the updated patient's clinical process?
Question 8: In your opinion, it would be appropriate the support of the new technologies to the daily provision of health care in your unit?

A statistical analysis of data was performed, through the results obtained. These quantitative results were normalized to realize a qualitative analysis of data, thus identifying patterns and trends. This analysis was chosen contrary to the Data Mining process, because the results sample is relatively low. Consequently, problem identification is immediate and direct.

The questionnaires were handled one by one through Tableau Software, a tool that allows you to organize and interpret data, easily answering qualitative questions and identifying patterns and trends.

The most Business Intelligence (BI) users use Data Visualization Tools to handle their raw data and transform them into useful and informative data. Nowadays, Tableau Software is one of the most widely used data analysis and visualization tools in the world. Compared to other tools such as QlikView and MicroStrategy Analytics, Tableau provides three versions (online, desktop and server version). Combined

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

with its high speed, this software is easy and User-friendly, and contains a highly competitive feature, the access with multiple platforms. Currently, its list of clients is very varied, with names like: Deloitte, eBay, Google, The New York Times, PayPal, Skype and many more (Reddy, 2015).

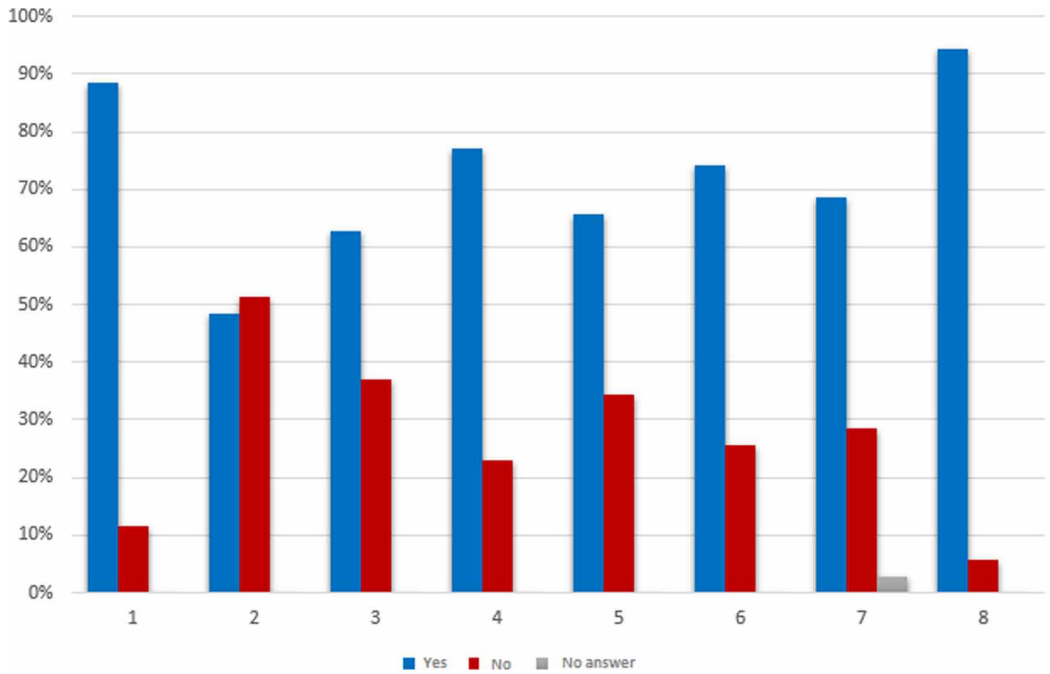
The data entered in the data source were of the nominal type. Thus, the value 0 was used to "No" answers and the value 1 to answers "Yes". To facilitate the analysis and understanding of the data, the results are presented as a percentage. The answers to the questionnaire can be found in Figure 4.

The interpretation of the results led to several of the initial motivations of the project being strengthened and reinforced. The idea that the support of the new technologies in the provision of healthcare in a hospital unit is fundamental. The need for integration and interoperability between systems, devices and health professionals is vital.

NURSING PRACTICE SUPPORT SYSTEM

The support system for nursing practice in an inpatient setting aims to help the health professional, in this case nurses, in real time. Nurses should obtain a more quick and effective access to all kinds of useful information about the patients and their tasks and actions at their health department. Besides, the application has a repository component that can be accessed at any time, called Patient History. Here each patient will have all their detailed information from the beginning of their hospitalization. The Patient's History was built according to the results obtained in the questionnaires mentioned above. It is very important for the process of nurse shifts. This feature ensures that no information is lost and is visible for all health professionals.

Figure 4. Result of the answers to the four questions in the questionnaire analyzed



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Characteristics

The main characteristics of the Nursing Practice Support System are described below:

1. Frequency of Data Update:
    - Minimum: Day to day;
    - Maximum: 5 to 5 minutes.
2. Usability:
    - Real-time updated patient information;
    - Simple and effective organization of information, making it easy to perceive and consult;
    - Alerts of future events, in particular, analyzes and examinations for executing or changing medications and types of diets, of each patient;
    - Easy location of inpatient in the health unit;
    - Availability of the daily tasks of nurses in their healthcare department;
3. Information Content:
    - Medication administration;
    - Upcoming Examinations, Analyzes, and Surgeries;
    - Diagnostics;
    - Type of Diets;
    - Location of patient in health unit;
    - Daily Nursing Tasks.

Data Sources

The information in the support system for nursing practice during hospitalization comes from different types of data sources:

- Electronic Nursing Record;
- Electronic Health Record;
- Agency for the Integration, Diffusion and Archiving of Clinical Information;
- Drugs System.

For this project to be possible, there was a need to gather information from several HIS. Consequently, it was developed a process capable of executing all the phases of extraction, transformation and loading (ETL) of the information in an autonomous way.

This new system of support to hospitalization has as main features the high interoperability existing between the different existing systems. Thus, the system avoids that during its tasks the health professional must manually access each one of these systems.

Multi-Agent System

As can be seen in Figure 5, eight agents were created with sequential and cyclic behaviors, and the first seven are responsible for completing each auxiliary table (Diet, Medication, Analyzes, Examinations, Tasks, surgeries). To avoid duplication of information, a truncate is done to each one of them before

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the information is entered. It was decided to create different agents for each auxiliary table to not compromise the performance of actions. Each table contains an attribute named "flag" with initial value 0. When the auxiliary tables are fulfilled with the new information (by the first seven agents), the flag of each one of them takes its initial value (0) and then the eighth agent updates the main table with that information, checking if the value of the flag is 0 (means is new information). In the end, he updates the flag of each auxiliary table to 1.

DISCUSSION

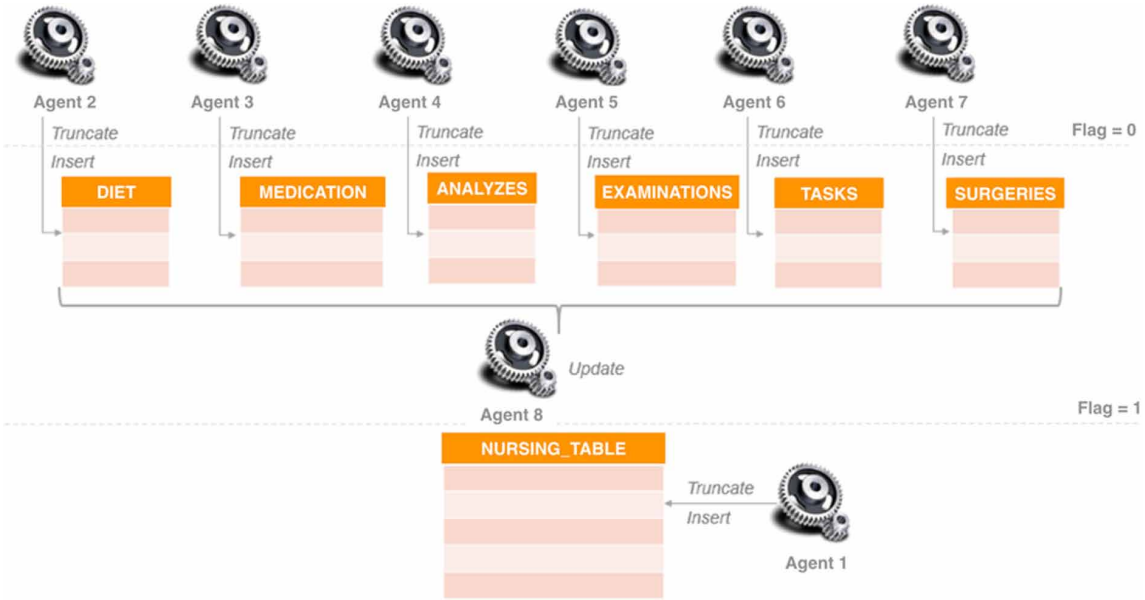
The support system for inpatient nursing practice is a project under development. It was previously discussed and approved by clinical leaders of the healthcare institution and later by health professionals who will be their future users. Through this new system, the information is getting available in real-time. For this, it is used autonomous agents to constantly execute ETL processes constantly. The update frequency depends on the refresh rate associated with each kind of data.

The fact that the system presents only the most relevant information, most succinctly and assertively as possible, confers to the application a high level of security and, consequently, a high confidence in the presented solution.

The major advantages of this project are:

- Improving quality of health record by the health professionals;
- Improving quality of the data, from its collection to its processing and validation;
- Increasing of the number of relevant information about the patient in the same place;
- Easing access to information (easy reading and interpretation).

Figure 5. Schematization of the multi-agent system developed



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The platform was developed respecting the principles of information automation, interoperability, usability, adaptability, integrality, distribution and competition. Therefore, the result was a web application easy to understand and use, with the ability to adapt to different operating systems and HISs.

The first implementation and testing phase of the prototype was carried out at the CHP hospital, which demonstrated the feasibility of the solution.

CONCLUSION

Increasingly, the use of new technologies for providing better healthcare is emerging. In this context, the proposed system aims to contribute to improving the patient's healthcare quality. Thus, the solution focuses on the patient and his well-being, reducing the possibility of the exposure to clinical error. The health professionals seem to be enthusiastic and confident with the solution. They have the awareness that the solution can be of great impact in their daily work.

The main motivations for this project regarding the results obtained in the questionnaires performed to health professionals were the decrease of the probability of clinical error or adverse events; and the creation of interoperability bridges between the several HIS holders of the different kinds of gathered information.

As future work, this application can be extended to other health departments of the hospital, suffering adaptability depending on the requisites of each service. Another improvement will be its portability to a mobile version, where health professionals will be able to consult it using a Smartphone or a Tablet with different operating systems. The system also will be a target of several studies of usability and technology acceptance. In the next few steps, the interface can undergo some improvements.

ACKNOWLEDGMENT

This work has been supported by Compete POCI-01-0145-FEDER-007043 and FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

REFERENCES

American Hospital Association. (2015). Why interoperability matters. Retrieved November 22, 2017, from http://www.aha.org/content/15/interoperabilitymatters.pdf

Benbasat, I., Goldstein, D., & Mead, M. (1987). The case research strategy in studies of information systems. *Management Information Systems Quarterly*, *1*(3), 220–247.

Berner, E. S., & Moss, J. (2005). Informatics challenges for the impending patient information explosion. *Journal of the American Medical Informatics Association*, *12*(6), 614–617. doi:10.1197/jamia.M1873 PMID:16049224

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cardoso, L., Marins, F., Portela, F., Santos, M., Abelha, A., & Machado, J. (2014). The Next Generation of Interoperability Agents in Healthcare. *International Journal of Environmental Research and Public Health*, *11*(5), 5349–5371. doi:10.3390/ijerph110505349 PMID:24840351

Coiera, E. (2003). *Guide to Health Informatics* (2nd ed.). London: Hodder Arnold. doi:10.1201/b13618

Donaldson, M. S., Corrigan, J. M., & Kohn, L. T. (2000). *To err is human: building a safer health system* (Vol. 6). National Academies Press.

Dove, J. T. (2005). The electronic health record—the time is now. *The American Heart Hospital Journal*, *3*(3), 193–200. doi:10.1111/j.1541-9215.2005.04389.x PMID:16106141

Duarte, J., Amorim, M., & Miranda, F. (2016). An Approach for the Semantic Interoperability of SNOMED: Improving Quality of. In *Applying business intelligence to clinical and healthcare organizations*.

Duarte, J., Portela, F., Abelha, A., Machado, J., Santos M., (2011). Electronic Health Record in Dermatology Service. In *Communications in Computer and Information Science*. Springer.

Duarte, J., Salazar, M., Quintas, C., Santos, M., Neves, J., Abelha, A., & Machado, J. (2010), Data Quality Evaluation of Electronic Health Records in the Hospital Admission Process. In *Proceedings of the 9th IEEE/ACIS International Conference on Computer and Information Science ICIS '10*. 10.1109/ICIS.2010.97

HIMSS. 2017, What is interoperability? Retrieved 24/04/2017 from http://www.himss.org/library/interoperability-standards/what-is-interoperability

Inmon, W. (1996). *Building the Data Warehouse*. John Wiley.

Khodambashi, S. (2013). Business Process Re-engineering Application in Healthcare in a Relation to Health Information Systems. *Procedia Technology*, *9*(2212), 949–957. doi:10.1016/j.protcy.2013.12.106

Kim, Y., & Park, H. A. (2011). Development and validation of detailed clinical models for nursing problems in perinatal care. *Applied Clinical Informatics*, *2*(2), 225–239. doi:10.4338/ACI-2011-01-RA-0007 PMID:23616872

Lang, E., Afilalo, M., Vandal, A. C., Boivin, J.-F., Xue, X., Colacone, A., ... Rosenthal, S. (2006). Impact of an electronic link between the emergency department and family physicians: A randomized controlled trial. *CMAJ: Canadian Medical Association Journal*, *174*(3), 313–318. doi:10.1503/cmaj.050698 PMID:16399880

Machado, J., Abelha, A., Neves, J., & Santos, M. (2006). Ambient Intelligence in Medicine. In *Proceedings of the Biomedical Circuits and Systems Conference, Healthcare Technology IEEE-Biocas '06*,, Imperial College, London, UK.

Machado, J., Alves, V., Abelha, A., & Neves, J. (2007). Ambient Intelligence via Multiagent Systems in Medical arena. *International Journal of Engineering Intelligent Systems, Special issue on Decision Support Systems*, *15*(3), 167-173.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Maria Salazar. (2013). Júlio Duarte, Rui Pereira, Filipe Portela, Manuel Santos, António Abelha and José Machado, Step towards Paper Free Hospital through Electronic Health Record. In *Advances in Information Systems and Technologies, Advances in Intelligent Systems and Computing* (Vol. 206). Springer.

Miranda, M., Salazar, M., Portela, F., Santos, M., Abelha, A., Neves, J., & Machado, J. (2012). Multiagent systems for hl7 interoperability services. *Procedia Technology*, *5*, 725–733. doi:10.1016/j.protcy.2012.09.080

Moreira, M. W., Rodrigues, J. J., Kumar, N., Al-Muhtadi, J., & Korotaev, V. (2017). Evolutionary radial basis function network for gestational diabetes data analytics. *Journal of Computational Science*. doi:10.1016/j.jocs.2017.07.015

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77. doi:10.2753/MIS0742-1222240302

Peixoto, H., Santos, M., Abelha, A., & Machado, J. (2012). Intelligence in Interoperability with AIDA. In *Proceedings of the 20th International Symposium on Methodologies for Intelligent Systems, Macau*, *LNCS* (Vol. 766). Springer.

Portela, F., Cabral, A., Abelha, A., Salazar, M., Quintas, C., Machado, J., . . . Santos, M. (2014). Knowledge Acquisition Process for Intelligent Decision Support in Critical Health Care. In Healthcare Administration: Concepts, Methodologies, Tools, and Applications. Hershey, PA: IGI Global.

Pozzebon, M., & Freitas, H. (1998). Pela aplicabilidade - com um maior rigor cientifíco - dos estudos de caso em sistemas de informação. *Revista de Administração Contemporânea*, *2*(2), 143–170. doi:10.1590/S1415-65551998000200009

Reason, J. (2000). Human error: Models and management. *The Western Journal of Medicine*, *172*(6), 393–396. doi:10.1136/ewjm.172.6.393 PMID:10854390

Reddy, V. (2015). Why is Tableau better than other Data Visualization Tools? *Linkedin*. Retrieved November 09, 2017, from http://www.linkedin.com/pulse/why-tableau-better-than-other-data-visualization-tools-vyshnavi-reddy/

Rocha, A., & Rocha, B. (2014). Adopting nursing health record standards. *Informatics for Health & Social Care*, *39*(1), 1–14. doi:10.3109/17538157.2013.827200 PMID:23978184

World Health Organization. (2015). European compendium of good practices in nursing and midwifery towards Health 2020 goals.

This research was previously published in the International Journal of Public Health Management and Ethics (IJPHME), 3(1); edited by Qiang (Shawn) Cheng; pages 26-37; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 47
An Analysis of Factors Affecting Postnatal Depression Intervention Adherence

Omobolanle Omisade
University of Portsmouth, UK

Alice Good
University of Portsmouth, UK

Tineke Fitch
University of Portsmouth, UK

Jim Briggs
University of Portsmouth, UK

ABSTRACT

Adjunct mobile support for postnatal depression could promote treatment adherence and long-term maintenance of behavioural change. The aim of this article is to establish the factors that determine adherence to postnatal depression intervention and support. Also, this article is intended to establish attitudes that women have towards postnatal depression intervention and support. Eighty-four women with a previous diagnosis of postnatal depression completed an online questionnaire on their previous use of postnatal depression intervention and factors inhibiting adherence, as well as attitudes towards the intervention. Results showed that adjunct support and combining multiple interventions would improve adherence. The provision of treatment guidance will also positively enhance treatment uptake and retention. Therefore, these factors should be considered for the development of theory-based adjunct mobile application for postnatal depression.

DOI: 10.4018/978-1-7998-1204-3.ch047

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Women are vulnerable to the affective mood disorder postnatal depression (PND), which can occur in the first three months after childbirth (Culjak & Spranca, 2006; O'Mahen et al., 2013). From the literature, PND represents the most frequent form of maternal morbidity following delivery (Cuijper, Brannmark, & Van Straten, 2008; Gibson, McKenzie-McHarg, Shakespeare, Price, Gray, 2009). A meta-analysis of 59 studies reports that 13% of women having babies suffer from PND with about 70,000 women experiencing PND in the United Kingdom every year (Caramlau, Barlow, Sembi, McKenzie-McHarg, & McCabe, 2011; Dennis, 2003; Evans, Donelle, Hume-Loveland, 2011; Glover, Onozawa, & Hodgkinson, 2002; Morrell, 2006). Like other episodes of depression, PND affects a postnatal woman's feelings about herself and her interpersonal relationships; and she may be functioning only minimally in her role as a mother (Beck, 2001). Additionally, PND can have serious consequences for the infant, which can include lower weight; impaired mental and emotional development; difficult temperament; poor self-regulation; low self-esteem; sudden infant death syndrome; or an overall higher frequency of hospital admissions and long-term behavioural problems (Forman, Videbech, Hedegaard, Salvig, & Secher, 2000; Glover et al., 2002; Grace, Evindar, & Stewart, 2003 Morrell, 2006; O'Mahen, 2013).

Antidepressant medication is usually the first treatment offered to treat the symptoms of PND. There is evidence to support its efficacy when used as prescribed, but it is often associated with side effects (Hou et al., 2013; Kaltenthaler et al., 2002). A study compared the effectiveness of antidepressants with placebo for the treatment of PND in an 8-week study (Yonkers, Lin, Howell, Heath, & Cohen, 2008). Women were randomised to either a placebo group or to take paroxetine antidepressant. Seventy women qualified for the study and only 31 completed between 7 and 8 weeks of treatment. While results indicated that there was a great improvement in the overall clinical severity for the paroxetine group compared with the controlled group, the study was restricted by high attrition rate and this complicated the interpretation of the findings. The high attrition rate in PND interventions remains a major problem that needs to be solved. One possible reason for high attrition for the use of antidepressants might be related to the fact that postnatal depressed mothers consistently indicate that they prefer therapy to antidepressants because of the concern over safety when breastfeeding (Hou et al., 2013; Ruaro, 2013). However, the authors cannot make a general conclusion.

Alternatively, psychotherapy is an effective treatment and support for PND. Milgrom et al. (2006) suggested that counselling informed by the principles of Cognitive Behavioural Therapy (CBT) (a form of psychotherapy) is effective when delivered by maternal and child health nurses. Interpersonal Therapy (IPT) has also been found to be effective for this population (Klier, Muzik, Rosenblum, & Lenz, 2001). However, the postnatal period still presents specific barriers to adequate improvement in depression, and a large number of women remain without treatment, do not adhere to the prescribed intervention or show only moderate improvement in depression (Boath, Bradley, Anthony, 2004; O'Mahen, 2013). More recently, computerised self-help approaches using the Internet have been developed to provide greater access and flexibility of receiving psychotherapy while requiring minimal input from a therapist (O'Mahen, 2013). The National Institute for Health and Care Excellence (NICE) has recommended the use of Computerized Cognitive Behavioural Therapy (CCBT) software for depression (NICE, 2006). CCBT has shown to be effective, with less time required in usual care and shorter therapist time needed while providing engagement and quality support (Churchill, et al., 2001; Gerhards, et al., 2010; Giridher, Wasilewska, Wong, & Rekhi, 2010; Graham, Franses, Kenwright, & Marks, 2000; Ingram, 2013; Kaltenthaler, Parry, Beverley, & Ferriter, 2008; Kaltenthaler et al., 2002; Marks, Shaw, & Parkin, 1998;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Merry et al., 2012). However, in clinical settings where women are only able to access CCBT from health centres, there is still the problem of long waiting lists for assessments and support (Andersson et al., 2005; Henshaw, 2004; Marks, Cavanagh, & Gega, 2007; O'Mahen et al., 2013; Proudfoot et al., 2004). The number of sessions of CCBT ranges from four to a maximum of 12 sessions and, in some cases, these sessions do not fit into a baby's variable and demanding schedule, resulting in poor adherence and inequality in access to appropriate treatment (Kaltenthaler et al., 2002; O'Mahen et al., 2013; Wan, Hu, Moore, & Ashford, 2008).

Research suggests that, despite effective treatment options, a large number of depressed women comply poorly with treatment or show only moderate improvement in depression (McCarthy & McMahon, 2008). In some cases, depressed women with other young children attributed poor adherence to struggles and cost associated with transportation or overwhelming responsibilities that can interfere with the ability to attend regularly scheduled appointments (De Graaf et al., 2009; Kaltenthaler et al., 2008; McCarthy & McMahon, 2008; Merry et al., 2012; Wan et al., 2008). Other barriers include stigma, long treatment sessions, sleep difficulties associated with infant sleep schedules and problems adjusting to and managing busy schedules of an infant balanced against other valued tasks (Kaltenthaler et al., 2008; O'Mahen et al., 2013).

The increase in the use of mobile phones and mobile applications may offer additional opportunities to provide support and may circumvent many of the difficulties that lead to poor adherence in PND. An adjunct mobile application could potentially be used to motivate postnatal depressed women to adhere to treatment and sustain treatment outcome in the long-term by providing guided support (Cuijpers, Donker, van Straten, Li, & Andersson, 2010; O'Mahen et al., 2013). It could offer the opportunity to provide just-in-time support and resources to those that have particular needs for flexibility (Kaltenthaler et al., 2008). A mobile application for PND would enable personalised interaction in a more practical and nonintrusive fashion as well as specialised treatment adaptation (O'Mahen et al., 2013). However, to develop an intervention that will increase adherence, there is a need to identify the factors that determine positive intentions to adhere to PND treatment. Furthermore, if you want to change the attitude towards a behaviour, you need to find out the beliefs people hold and the consequences of their attitude. A knowledge of these will also help to adapt the right kinds of beliefs, and, thus, a change in behaviour (Ajzen, 2002).

Studies have shown that, for optimum effectiveness, an intervention should have a sound theoretical basis. According to the Theory of Planned Behaviour (TPB), human action is influenced by three major constructs: behavioural beliefs/attitude (favourable or unfavourable), subjective norm (pressure to perform or not perform the behaviour) and control belief (perceived capacity to perform the behaviour) (Ajzen, 2002). They are all conceptualised as directly related to intentions towards behaviour. The targets of this paper are the behavioural beliefs and attitudes women have towards adhering to PND treatments as well as the control beliefs. Attitude towards a behaviour is assumed to be a function of beliefs about the behaviour's likely consequences, together with the valence attached to those consequences (Ajzen, 2011). It is assumed that a knowledge of women's beliefs and attitudes towards intervention outcome could predict their behaviour towards treatment adherence and their intention to use a support application. Fishbein advocates identifying salient beliefs from the intended population and developing suitable and appropriate materials based on the elicited beliefs (Fishbein, von Haeften, & Appleyard, 2001). Furthermore, a person may hold a positive attitude concerning a behaviour, but not always carry out their intentions towards it (Ajzen, 2011). In this context, a postnatal depressed woman might have positive intentions towards treatment, but there are many reasons as to why she does not adhere to it, and this can be classified as

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the degree of control that the woman has over behavioural performance. If a depressed woman has the capacity, skills, time, resources or whatever it takes to perform the adherence behaviour, then she will be able to carry out her intention, but not if some of these are lacking. However, women with PND are faced with overwhelming responsibilities and undergo some changes after childbirth that could affect their adherence behaviour (De Graaf et al., 2009; Kaltenthaler et al., 2002). There is, therefore, need to understand women's beliefs and attitudes toward interventions, control beliefs and how they form their intentions to engage, perform or not the adherence behaviour.

Identifying women's beliefs and attitudes towards PND interventions could potentially benefit practitioners who intend to change or motivate the adherence behaviour. In this context, this knowledge will inform the development of a theory-based adjunct mobile application to support the wellbeing of women with PND. It will also help to decide on the factors and beliefs that should be supported or discouraged by a mobile application for PND. The authors envisage that an adjunct mobile application would promote the self-monitoring and social support needed to motivate treatment adherence and long-term maintenance and control over adherence behaviour.

Study Aims and Objectives

This study aims to investigate the TPB constructs, behavioural and control beliefs to establish some of the factors that lead to treatment adherence and sustained outcome. The objectives are to identify the obstacles to the performance of adherence behaviour that should be discouraged within the mobile application. For women who are already motivated, the authors will also identify factors that could facilitate them to implement their intentions to adhere to treatment. This study will be more related to the type of treatment used and its effectiveness, as opposed to relaying past experiences of PND specifically.

METHOD

Design

The study used an online survey to collect data from participants. Ethical approval was obtained from the author's University Ethics Committee (ID number - BO1/OM1)

Recruitment

Participants were recruited by advertising on UK-based social media platforms, sending email invitations to women who are members of parenting support groups and fertility forums. This took place from February to May 2016. Moderators of the groups were consulted concerning the research, prior to actively targeting potential participants. Once approval had been granted, recruitment involved voluntary participation of women with previous diagnosis of PND.

Interested participants provided informed consent prior to participating and then accessed information describing the aims and objectives of the study and eligibility. It was recognised there may be a possibility that participants might have found that some of the study questions triggered distressing memories; therefore, the authors designed the questionnaire to ensure that questions related more to the type of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

therapy used and its effectiveness, as opposed to relaying past experiences of PND specifically. In the event they became distressed, the study also provided a set of resources that could support women. The authors did not ask any questions relating to children being at risk or other likely distressing memories.

Participants were provided with sufficient details of the study at an appropriate level of understanding and were informed that they had the right to withdraw at any time during the course of completing the questionnaire. Given the possibility, albeit minimal, that participants may incur psychological stress in relation to carrying out the survey, information was provided at the end of the survey advising women where they could seek support. The survey was designed to exclude any women who had a current or recent diagnosis of PND.

Instrument

Ajzen's (2002) instructions for constructing a TPB questionnaire were adopted to develop items used to collect data on treatment adherence behaviour. This study is simply measuring the attitude towards PND treatment adherence and the perceived control beliefs; therefore, this was measured by eight items with possible responses ranging from strongly disagree to strongly agree. See Figure 1 for how the construct was operationalised in the current study. The questionnaire included questions about the types of intervention participants have used in the past, the effectiveness of intervention on wellbeing, the effect of intervention on symptoms, factors that might cause non-adherence and demographics. See appendices for questionnaire items.

For the purpose of this study, eight PND interventions are considered in order to identify participant's attitude towards treatment adherence:

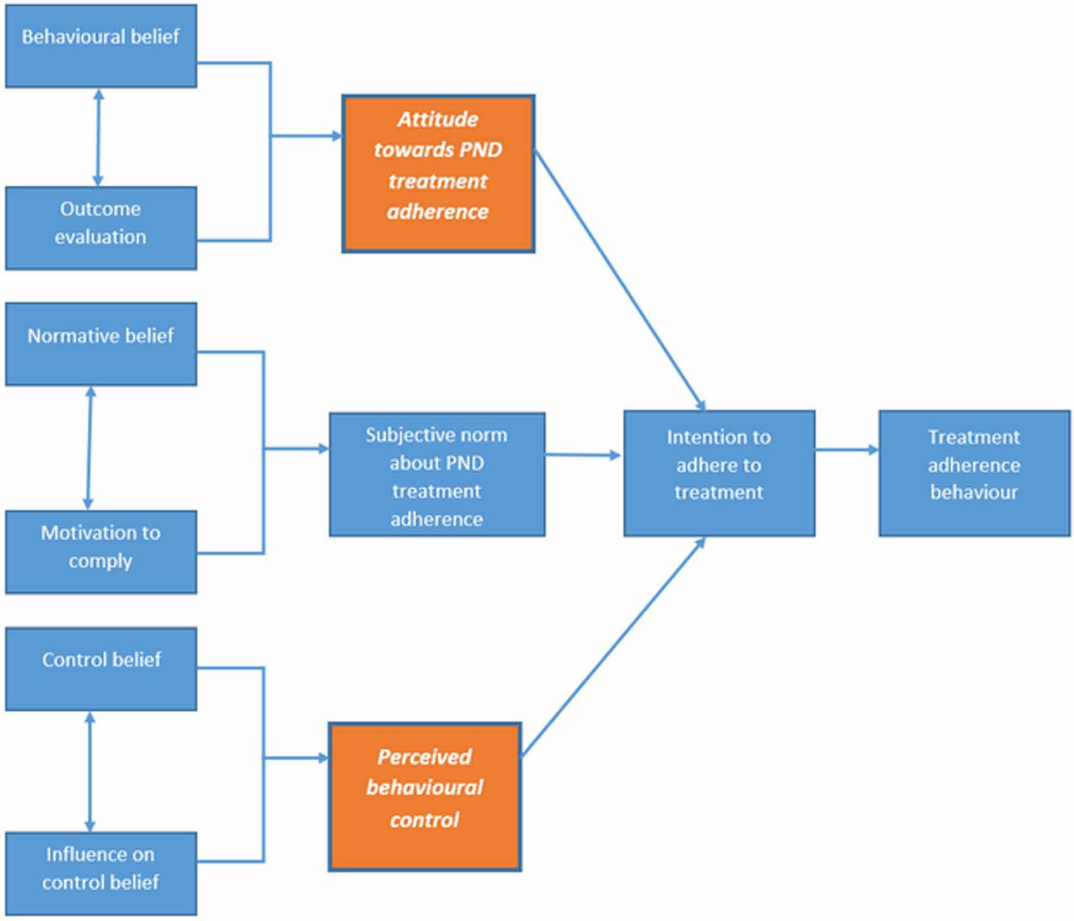
- Cognitive Behavioural Therapy (CBT)
- Medication (Med)
- Monitoring and Assistance (MA)
- Self-Help from peer group (SHP)
- Psycho-education
- Wellbeing therapy
- Interpersonal therapy
- Problem solving therapy

The natures of these interventions are described in the following studies: Klier et al. (2001), Cooper (2003), Milgrom, Ericksen, McCarthy, & Gemmill (2006), Pampallona et al. (2002 and Yonker et al. (2008). This study classified symptoms as:

- Isolation
- Poor communication
- Prioritising demand
- Relationships
- Stress management
- Negative thoughts
- Self-confidence

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Adapted Theory of Planned Behaviour Construct for predicting treatment adherence behavior (Ajzen, 2002)



Participants

Potential participants were screened for eligibility using multiple questions in the questionnaire. In order to determine whether a participant is currently ill or well, the authors asked the question "Do you presently suffer from postnatal depression". Eligible participants were women with previous diagnosis of PND, yet who had not suffered any symptoms of PND in the last two years prior to recruitment. This study group was used because it places us in a position to explore women's view about the short and long-term effectiveness of the intervention on their recovery from PND. Men, women who currently suffer from PND, have experienced PND in the last two years, or have not received treatment in the past were excluded from the study.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sample

This study estimated a sample size of 382 postnatal mothers to achieve effect size. This is based on a population of about 70,000 women suffering from PND in the UK (Caramlau et al., 2011; Dennis, 2003; Evans et al., 2011; Glover et al., 2002), with 5% margin of error and confidence level of 95%. 460 women were invited and, in total, 181 participants started the survey. After screening for depression, 97 were not eligible for the study. Of those, 55 (30.4%) were excluded because they were current sufferers of PND, 11 because they had never suffered from PND and 23 (12.7%) because they had not received any treatment for PND. A further eight respondents that did not complete any further questions had to be excluded due to the lack of further responses to analyse. The final group of participants comprised 84 (46.4%) postnatal women (mean age = 32 years, range 27-41). It is important to note that not all participants responded to all questions asked. Table 1 presents further details on the number of included and excluded participants.

Data Analysis

Throughout this study, the widely used IBM SPSS statistics package was used to analyse quantitative data (Brace, Kemp, & Snelgar, 2012). For this study, an outcome is statistically significant if the significance level "p" is less than 0.05. This study explored the differences in scores of a non-parametric dependent variable between three or more groups of a single independent variable. Therefore, the Kruskal-Wallis test, which is a non-parametric equivalent of the between subject analysis of variance (ANOVA) test, was employed to evaluate significant differences. The Kruskal-Wallis test only tells us that there is a difference, but does not tell us where the differences are according to each pair of analyses (Brace, et al., 2012). Therefore, an additional Mann-Whitney U test was used to locate the differences, one for each pair of groups.

Furthermore, this study measured the degree of relationship and association between symptoms and adherence variables. However, it is important to note that correlation does not imply causation, but will provide a measure of strength and direction of relationship. The Pearson correlation was used for normal data. For non-parametric data, Kendall's Tau-b test of correlation was used to predict the relationship between variables. The study assessed relationships in terms of the positive or negative correlation coefficient (r). The correlation coefficient describes the strength of the association between

Table 1. Number of included and excluded women

	Excluded women	**Included women**	**Percentage (%)**
Women not suffering from PND	11		6.1
Women not treated for PND	23		12.7
Women excluded due to the lack of further responses to analyse	8		4.4
Current sufferers	55		30.4
Treated and not currently suffering		84	46.4
Total	97	84	100.0

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

two variables. It provides a direct representation of effect size. Using Cohen's (1988) conventions for categorising effects sizes, the study described the strength and size of the relationship for particular values of correlation coefficient.

RESULTS

Depressed Women Without Treatment

Our survey revealed that, despite effective treatment options, some women remain without treatment. Although those women who had PND and remained without treatment were excluded from the main part of this study (see Table 1), our sample showed that 12.7% (23/181) of respondents fell into this category. Our findings agree with the literature that suggests that, despite effective intervention options, postnatal depressed women remain without treatment (Kaltenthaler et al., 2002; O'Mahen, 2013; Wan et al., 2008).

The Use of Multiple Interventions

Our results showed that 81% of depressed women rely on combinations of treatment strategies, with medication being the most common intervention that was combined. Figure 2 presents the frequency of the number of women who used multiple interventions. Table 2 shows that women found one intervention more effective than the other.

Can Combine Multiple Intervention Aid Symptoms

Further tests were carried out to identify women's attitudes towards the benefits and purpose of using multiple interventions. The Kruskal-Wallis test was used to identify differences between using one intervention as opposed to combining multiple interventions for PND symptoms. The results show that there is no significant difference between using one, two, three, four, five or six interventions on PND symptoms. Non-significance was apparent across all the symptom factors except relationship: $\chi 2\ (5) = 12.942$, $p = .024$, see Table 3.

Is There Similarity in Intervention Effects on Symptoms

All women in the study reported having all the symptoms used in this study. This study assumes that several interventions will have the same effect on symptoms. A statistical test was used to identify the belief of how the effects of intervention on PND symptoms differ. The authors identified significant differences between intervention groups about isolation, communication and managing negative thoughts. The result for prioritising demands, relationship, self-confidence, and stress management that were not statistically significant. A subsequent Mann-Witney U-test showed that these differences were particularly apparent between CBT and SHP ($z = -2.318$, $p = 0.02$), MED and SHP ($z = -2.540$, $p = 0.01$). Concerning scores on isolation and for communication, differences were found between the MED and SHP categories ($z = -2.863$, $p < 0.05$) and between managing negative thought CBT and SHP ($z = -2.451$, $p = 0.02$). Although, the results were apparently between neutral and disagree, it shows that various interventions have different effects on symptoms.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Frequency of number women who used multiple interventions

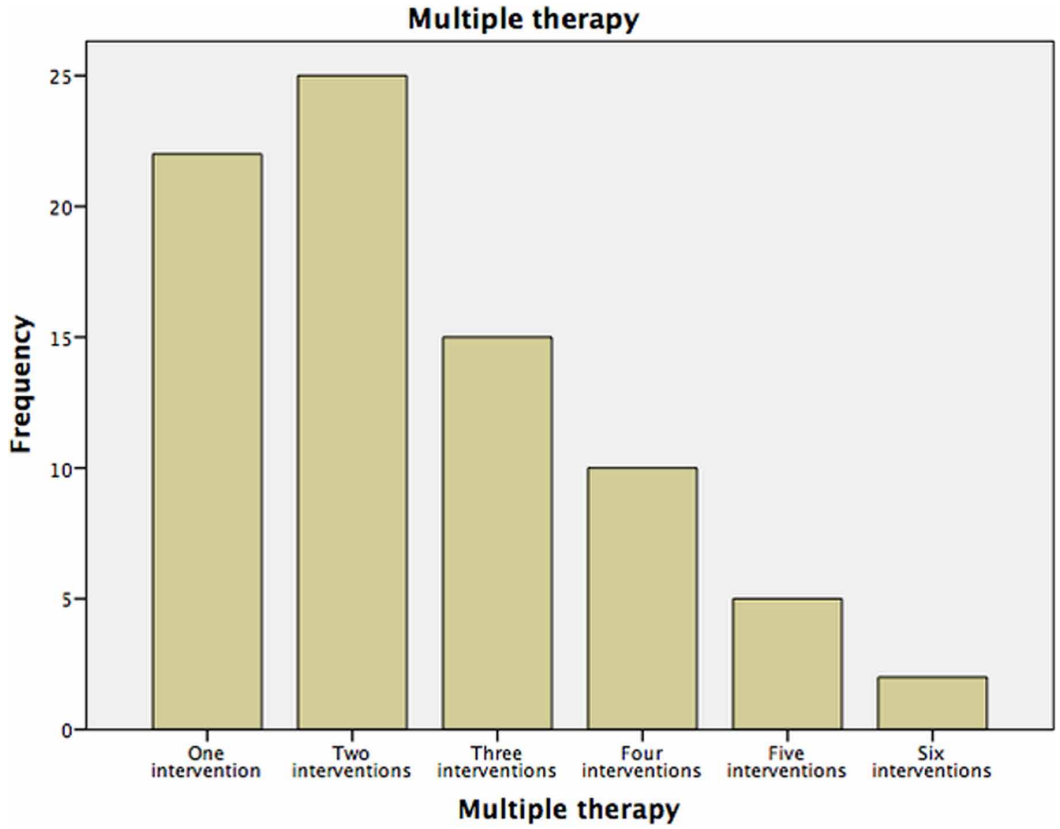


Table 2. Frequency of participants' most effective intervention and percentage score of women that have used multiple interventions

Intervention	Number of women who used intervention	Number of women for whom this was the most effective intervention	Number of women for whom this was one of a set of multiple interventions
Cognitive Behavioural Therapy	29	12	12
Monitoring and assistance (GP, Nurse or Therapist)	37	6	6
Medication	67	40	27
Psycho-education and information	3	1	
Self-help from peer group	26	10	7
Interpersonal Therapy	7	3	5
Wellbeing Therapy	5	5	5
Problem Solving Therapy	2	2	2
Total	176	79	64

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Effects of using multiple interventions to aid symptoms

Symptoms factors	Kruskal-Wallis test		
	H	df	P-value
Isolation	2.794	0.73	0.73
Communication	2.042	0.84	0.84
Prioritise demands	2.995	0.71	0.71
Relationship	12.942	5.00	0.02
Stress management	6.392	5.00	0.27
Manage negative thoughts	2.626	5.00	0.76
Self confidence	6.399	5.00	0.27

Association Between Non- Adherence and PND Symptoms

This study assumes that there is a relationship between non-adherence factors and the effect of PND intervention on symptoms. To evaluate this assumption, the authors considered poor adherence control factors available from the literature to be the cost of receiving intervention; the duration of appointments; waiting times; and location convenience. All women in the study reported facing the control factors used in this study. The results in Table 4 show that there was a weak and non-significant correlation between most of these. Although the correlation between the location and stress management was significant, it is a small effect (r= 249, p= .01).

Furthermore, we identified how women followed treatment instructions as prescribed by their practitioners. Research suggests that half of those women for whom treatment regimens were prescribed do not follow them as directed - this is termed as non-adherence (World Health Organisation, 2003). The researchers asked women how often they followed their treatment instructions. Table 5 shows that 63% (mean = 3.77; sd = .89) of the final group of the women followed the instructions as directed by their practitioners. However, there was a weak and non-significant relationship between the frequency of following instructions and the effect on symptoms (see Table 6).

Table 4. Association between PND symptoms and factors that might affect adherence to PND treatment (control factors)

Symptoms factors	Kendall's Tau_b correlation							
	Cost factor		Duration factors		Waiting times factor		Location factor	
	r	p	r	p	r	p	r	p
Isolation	.151	.12	.068	.49	.013	.90	.019	.85
Communication	-.070	.45	.004	.97	.108	.27	.011	.91
Prioritise demands	.059	.54	.172	.08	.057	.57	.010	.91
Relationship	.015	.88	-.096	.32	.031	.76	.124	.20
Stress management	.023	.81	-.047	.63	-.113	.26	.249	.01
Manage negative thoughts	.006	.95	-.169	.09	-.088	.39	.109	.27
Self confidence	.068	.48	.050	.61	.001	.99	.059	.54

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Number of women that follow instructions as directed by practitioner

	Frequency	Percentage (%)
Never	4	6.1
Sometimes	20	30.3
Regularly	42	63.6
Total	66	100.0

Table 6. Summary of relationship between following instructions and symptom control factors

Symptoms factors	Kendall's tau_b correlation		
	r	P-value	Variance
Isolation	-.033	0.76	.729
Communication	.096	0.37	.921
Prioritise demands	-.113	0.29	.127
Relationship	-.021	0.85	.289
Stress management	.005	0.96	.25
Manage negative thoughts	-.066	0.55	.291
Self confidence	.081	0.45	.656

The Soothing of One PND Symptom Might Alleviate the Effect of Other Symptoms

The Pearson correlation coefficient was used to examine the degree of association between symptom factors. This test was used because there was a normal distribution of data. Prioritising demands was found to be significantly and positively correlated with communication ($r\ (n=79) = 0.464$, $p < 0.001$); relationship ($r(n=79) = 0.434$, $p < 0.001$); stress management ($r(n=79) = 0.408$, $p < 0.001$); managing negative thoughts ($r(n=79) = 0.329$, $p < 0.001$) and self-confidence ($r(n=79) = 0.474$, $p < 0.001$). These indicate that depressed women with a higher level of prioritising their demands tended to have a good relationship with people, have reduced stress, are able to manage negative thoughts and have faith in themselves. However, isolation was found to have a very weak relationship with prioritising demands ($r(n=79) = 0.164$, $p = 0.148$), communication ($r(n=79) = 0.208$, $p = 0.065$), and relationship ($r(n=79) = 0.177$, $p = 0.119$) (see Table 7).

DISCUSSION

This study documents the beliefs and factors that might determine adherence to PND treatment, and ensure increased sustained outcome. Firstly, a favourable attitude towards combining multiple interventions influences the actual adherence behaviour. There are several possible reasons why depressed women combine interventions, one being seeking support for normalisation of their mothering (O'Mahen et al., 2013). Previous research has shown that many women begin to incorporate alternate plans after

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. Association between symptoms considered in this study

Symptom	2	3	4	5	6	7
1. Communication	.208	.464**	.438**	.335**	.301**	.493**
2. Isolation	-	.164	.177	.252*	.325**	.308**
3. Prioritising demands		-	.434**	.408**	.329**	.474**
4. Relationships			-	.520**	.390**	.550**
5. Managing stress				-	.777**	.557**
6. Managing negative thoughts					-	.512**
7. Self confidence						-

the initial core sessions of their initial intervention (Christensen, Griffiths, & Farrer, 2009; O'Mahen et al., 2013). A poor patient-doctor relationship might cause some women to be inconsistent in attending treatment sessions, thus finding themselves using alternative treatments. In previous research, women have expressed negative experiences of seeking help, and this is particularly likely to cause them to seek additional treatments. Long waiting lists for treatments might be a barrier; therefore, they then seek additional support during waiting times. In some cases, treatments are not equivalently distributed geographically (Goodman 2009; Payne & Myhr, 2010). Finally, perceived stigma can also lead to women seeking additional support that will provide the level of anonymity they require (Dennis & Chung-Lee, 2006; Goodman, 2009; O'Mahen, et al., 2013). Although our results show that combining multiple interventions does not have any significant effect on symptoms, the authors identified, however, that different interventions have different levels of effect on PND symptoms, therefore combining multiple interventions should be increased. This is further supported by research that suggested that to improve adherence, medication should be combined with other support strategies such automatic reminders, reminder packaging, medication boxes, device aids, counselling, telephone support, patient education etc., (Desteghe et al., 2017).

In this present study, women's belief is that the soothing of one symptom factor alleviates the effects of other symptoms. The authors identified a weak relationship between symptoms; this could mean that women are encouraged to adhere or not when the effect of treatment on one or more symptoms is obvious. This is similar to a study that revealed that it was only after women got into a supportive treatment relationship and improved their communication that they felt able to disclose their distress to others (McCarthy & McMahon, 2008). However, further study may be needed to highlight the factors that suggest the relationship between PND symptoms and the effect they have on each other.

Another strength of this study is that treatment cost, duration, location and waiting time, identified as control factors that inhibit adherence to PND treatment, are not related to increased adherence, but not in all cases. One possible explanation is that these factors should not totally hinder women from adhering to prescribed treatments. However, the location where intervention is received is seen to have a relationship with stress management; therefore, it should be recognised that there may be a possibility that depressed women become distressed by the location of receiving treatment. This could be due to transportation difficulties faced by some women. In regard to location, studies have documented the inability of women to attend regularly scheduled appointments (Goodman, 2009; O'Mahen et al., 2014). Additionally, this might mean that an adjunct support should be provided as a control measure in the event they should become distressed by the treatment location. It is also useful to provide guidance as

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

moderator of intention-behaviour and, thus, this might enhance treatment uptake and retention (Cuijpers et al., 2010; O'Mahen et al., 2013). This compares favourably with other research suggestions that only women who get a supportive treatment are able to show sustained outcome (Cuijpers et al., 2008; O'Mahen et al., 2013; McCarthy & McMahon, 2008). Furthermore, following instructions as prescribed by practitioners did not have an effect on PND symptoms. The underlying causes as to why following instructions did not have effect on symptom are still unclear. Therefore, the authors are careful in making their conclusions on this finding.

Implications

This study identified women's attitudes towards postnatal depression interventions and the factors that could motivate improved adherence. A knowledge of these factors will allow the development of a set of guidelines which can inform the development of an adjunct mobile application for postnatal depression and future improvement research. This will also facilitate the development of a mobile application that is tailored towards positive treatment adherence attitudes. Furthermore, it will potentially support future applications for a larger study to undertake a full trial of mobile applications for PND.

Limitations

Limitations of the current study are largely related to the fact that not all participants provided us with a complete data set. Although all participants satisfied the eligibility criteria, the women who completed the study are best considered a convenient sample that may not be generalised to all postnatal depressed mothers. Finally, it is possible that the use of only two TPB constructs (behavioural and control beliefs) might have influenced the ability to identify more factors that facilitate or inhibit adherence behaviour.

CONCLUSION

This study has identified the TPB variables that might predict the intention to adherence to PND treatments. Firstly, increasing access to adjunct support for PND has the potential to influence adherence positively, and long-term sustained treatment outcome. In particular, providing an intervention that combines support features of several interventions is recommended for increased adherence. This intervention could also provide the flexibility of treatment adaptation. Additionally, it may circumvent many of the barriers to treatment delivery and motivate women to self-manage their wellbeing as well as give them access to supplementary wellbeing support.

It is essential that the option to combine multiple interventions should also be balanced, this could mean that it should not be encouraged or discouraged. This is because combining multiple interventions does not have any significant effect on symptoms; however, different interventions have different levels of effect on PND symptoms. Therefore, the authors recommend a platform that will facilitate using additional support to prescribed treatments for PND. This is in line with a growing body of research demonstrating that the provision of additional support to prescribed treatment could enhance achieving the desired outcome and sustained effect (Broom, Ladley, Rhyne, & Halloran, 2015). Adjunct treatment needs to remain available to women receiving intervention, both during and after the intervention has ended (Turner, Chew-Graham, & Folkes, 2010). The effect of adjunct support was revealed in a

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

study that recorded high compliance (Rojas et al., 2007). The multicomponent intervention used in that study was supported by regular phone calls from trained mental health workers, and depressed women were constantly reminded in sessions about the need to take medications as prescribed (Rojas, et al., 2007). The authors also recommend that, rather than providing women with a single intervention for all PND symptoms, there should be a multiple-choice option that is open to adapting support to depressed women's specific needs. This should hopefully give them the opportunity to have control over their desired intervention outcome.

Furthermore, this study demonstrated that a greater guidance tool designed to make treatment effect on symptoms noticeable could be beneficial and help control adherence behaviour. A structured support plan is viewed as a behavioural control tool, critical for achieving increased treatment understanding, self-management and adherence.

Given these promising findings, there is a need to develop and evaluate methods of promoting adherence and sustained treatment outcome in the PND population. Mobile Health is growing as a method to improve adherence and health outcomes. An adjunct mobile application is one method that may hold the promise of doing such. Developing a suitable and appropriate mobile application is desirable, given its potential to increase adherence to the behaviour being promoted (Desteghe et al., 2017). This tool could support individuals with treatment/work already offered by practitioners, thus allowing sessions to be offered at a reduced frequency, duration and at a cheaper cost and convenient location. In order to develop a mobile health intervention that will increase adherence, using a theory based approach, this study sought to understand women's perspective on obstacles to the performance of adherence behaviour that should be discouraged within the mobile application as well as factors that could facilitate them to implement their intentions to adhere to treatment. Based on our findings, we recommend that the development of a PND adherence mobile application should include features designed to:

- Help manage and monitor the use of multiple interventions.
- Keep a record and manage the impact of symptoms so that changes are noticeable.
- Keep records of prescription instructions and provide positive prompts to users for better communication.
- Include prompts and reminders as a means of self-guidance.
- Help users stay in contact with healthcare providers, family and friends when further guidance or help is needed.

Additionally, research has shown that self-incentives are key motivators and successful in changing health behaviours, adherence to daily care and improve health outcome. As a result of this insight we recommend that the mobile app should also include a feature that is:

- Related to self-incentives and will make the environment enjoyable and stimulating for users.

The current findings highlight beliefs about PND interventions that are necessary for developing adherence behaviour change intervention for this population. These findings will inform the design of an adjunct mobile application for PND. However, there is a need to establish the suitability and appropriateness of a mobile application. The mobile application should be evaluated for suitability and appropriateness to enhance the likelihood that the tool can effectively change behaviour. Further research

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

is also needed to determine the acceptance of such a mobile application by both women with PND and practitioners. Such research should allow us to establish the requirements under which a mobile application for PND may be offered as effective and accessible support.

REFERENCES

Ajzen, I. (2002). Perceived behavioural control, self-efficacy, locus of control and the theory of planned behaviour. *Journal of Applied Social Psychology*, *32*(4), 665–683. doi:10.1111/j.1559-1816.2002.tb00236.x

Ajzen, I. (2011). Behavioral interventions: Design and evaluation guided by the theory of planned behavior. In M. M. Mark, S. I. Donaldson, & B. C. Campbell (Eds.), *Social psychology for program and policy evaluation New York* (pp. 74–100). Guilford.

Andersson, G., Bergstrom, J., Hollandare, F., Carbring, P., Viktor, K., & Ekselius, L. (2005). Internet-based self help for depression. *The British Journal of Psychiatry*, *187*(5), 456–461. doi:10.1192/bjp.187.5.456 PMID:16260822

Beck, C. T. (2001). Revision of the Postpartum Depression Predictors Inventory. *JOGNN principles and practice*, *31*(4), 394-402.

Boath, E., Bradley, E., Anthony, P. (2004) Users' views of two alternative approaches to the treatment of postnatal depression. *Journal of Reproductive and Infant Psychology*, *22*(1), 13-24.

Brace, N., Kemp, R. and Snelgar, R.(2012). *SPSS for Psychologist* (5th ed.). Basingstoke, UK: Macmillan Publishers.

Broom, M. A., Ladley, A. S., Rhyne, E. A., & Halloran, D. A. (2015). Feasibility and perception of using text messages as an adjunct therapy for low-income, minority mothers with postnatal depression. *JMIR Mental Health*, *2*(1), e4. doi:10.2196/mental.4074 PMID:26543910

Caramlau, I., Barlow, J., Sembi, S., McKenzie-McHarg, K., & McCabe, C. (2011). Mums 4 Mums: Structured telephone peer-support for women experiencing PND. Pilot and exploratory RCT of its clinical and cost effectiveness. *Trials*, *12*, 88. doi:10.1186/1745-6215-12-88 PMID:21439042

Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in internet interventions for anxiety and depression. *Journal of Medical Internet Research*, *11*(2), e13. doi:10.2196/jmir.1194 PMID:19403466

Churchill, R., Hunot, V., Corney, R., Knapp, M., McGuire, H., Tyleem, A., & Wessely, S. (2001). A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment*, *5*, 1–173. PMID:12387733

Cooper, J. E. (2003). Detection and Management of Psychiatric Disorders in Primary Care. *The British Journal of Psychiatry*, *182*(1), 1–4. doi:10.1192/bjp.182.1.1 PMID:12509308

Cuijper, P., Brannmark, J. G., & Van Straten, A. (2008). Psychological Treatment of Postpartum Depression: A meta-analysis. *Journal of Clinical Psychology*, *64*(1), 113–118. PMID:18161036

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cuijpers, P., Donker, T., van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine*, *40*(12), 1943–1957. doi:10.1017/S0033291710000772 PMID:20406528

Culjak, G., & Spranca, M. (2006). Internet Improves Health Outcome in Depression. In *Proceedings of the Hawaii International Conference on System Sciences*.

De Graaf, L. E., Gerhards, S. A. H., Arntz, A., Riper, H., Metsemakers, J. F. M., Evers, S. M., ... Huibers, M. J. (2009). Clinical effectiveness of online computerised cognitive–behavioural therapy without support for depression in primary care: Randomised trial. *The British Journal of Psychiatry*, *195*(1), 73–80. doi:10.1192/bjp.bp.108.054429 PMID:19567900

Dennis, C. (2003). The effect of peer support on postpartum depression: A pilot randomized controlled trial. *Canadian Journal of Psychiatry*, *48*(2), 115–125. doi:10.1177/070674370304800209 PMID:12655910

Dennis, C.-L., & Chung-Lee, L. (2006). Postpartum depression help-seeking barriers and maternal treatment preferences: A qualitative systematic review. *Birth (Berkeley, Calif.)*, *33*(4), 323–331. doi:10.1111/j.1523-536X.2006.00130.x PMID:17150072

Desteghe, L., Vijgen, J., Koopman, P., Dilling-Boer, D., Schurmans, J., Dendale, P., & Heidbuchel, H. (2017). The Health Buddies App as a Novel Tool to Improve Adherence and Knowledge in Atrial Fibrillation Patients: A Pilot Study. *JMIR mHealth and uHealth*, *5*(7), e98. doi:10.2196/mhealth.7420 PMID:28724512

Evans, M., Donelle, L., & Hume-Loveland, L. (2011). Social support and online postpartum depression discussion groups: A content analysis. *Patient Education and Counseling*, *87*(3), 405–410. doi:10.1016/j.pec.2011.09.011 PMID:22019021

Fishbein, M., von Haeften, I., & Appleyard, J. (2001). The role of theory in developing effective interventions: Implications from Project SAFER. *Psychology Health and Medicine*, *6*(2), 223–238. doi:10.1080/13548500123176

Forman, N. D. (2000). Postpartum Depression: Identification of women at risk. *British Journal of Obstetrics and Gynaecology*, *107*(10), 1210–1217. doi:10.1111/j.1471-0528.2000.tb11609.x PMID:11028570

Gerhards, S. A. H., de Graaf, L. E., Jacobs, L. E., Severens, L. J., Huibers, M. J. H., Arntz, A., ... Evers, S. M. (2010). Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: Randomised trial. *The British Journal of Psychiatry*, *196*(4), 310–318. doi:10.1192/bjp.bp.109.065748 PMID:20357309

Gibson, J., McKenzie-McHarg, K., Shakespeare, J., Price, J., & Gray, R. (2009). A systematic review of studies validating the Edinburgh PND Scale in antepartum and postpartum women. *Acta Psychiatrica Scandinavica*, *119*, 350–364. doi:10.1111/j.1600-0447.2009.01363.x PMID:19298573

Giridher, I., Wasilewska, A., Wong, J. L., & Rekhi, K. S. (2010). Global mobile application for monitoring health. In *Proceedings of the 2010 International Multiconference on Computer Science and Information Technology* (pp. 855-859).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Glover, V., Onozawa, K., & Hodgkinson, A. (2002). Benefits of infant massage for mothers with PND. *Seminars in Neonatology*, *7*, 495–500. doi:10.1053iny.2002.0154 PMID:12614602

Goodman, J. H. (2009). Women's attitudes, preferences and perceived barriers to treatment for perinatal depression. *Birth (Berkeley, Calif.)*, *36*(1), 60–69. doi:10.1111/j.1523-536X.2008.00296.x PMID:19278385

Grace, S. L., Evindar, A., & Stewart, D. E. (2003). The effect of postpartum depression on child cognitive development and behavior: A review and critical analysis of the literature. *Archives of Women's Mental Health*, *6*(4), 263–274. doi:10.100700737-003-0024-6 PMID:14628179

Graham, C., Franses, A., Kenwright, M., & Marks, I. (2000). Psychotherapy by Computer: A postal survey of responders to a teletext article. *Psychiatric Bulletin*, *24*(9), 331–332. doi:10.1192/pb.24.9.331

Henshaw, C. A. (2004). What do women think about treatments for PND. *Clinical Effectiveness in Nursing*, *8*, 170–175. doi:10.1016/j.cein.2005.03.003

Hou, Y., Hu, P., Zhang, Y., Lu, Q., Wang, D., Yin, L., ... Zou, X. (2013). Cognitive behavioral therapy in combination with systemic family therapy improves mild to moderate postpartum depression. *Revista Brasileira de Psiquiatria (Sao Paulo, Brazil)*, *36*(1), 47–52. doi:10.1590/1516-4446-2013-1170 PMID:24604461

Ingram, D., & Moreno, M. (2013). A computerized self-help intervention is as effective as face-face counselling for adolescents seeking help for depression. *The Journal of Pediatrics*, *161*(5), 967–968. doi:10.1016/j.jpeds.2012.08.050 PMID:23095695

Kaltenthaler, E., Parry, G., Beverley, C., & Ferriter, M. (2008). Computerised cognitive-behavioural therapy for depression: Systematic review. *The British Journal of Psychiatry*, *193*(3), 181–184. doi:10.1192/bjp.bp.106.025981 PMID:18757972

Kaltenthaler, E., Shackley, K., Stevens, K., Beverly, C., & Parry, G. (2002). A systematic review and economic evaluation of computerized cognitive behaviour therapy for depression and anxiety. *Health Technology Assessment*, *6*(22). doi:10.3310/hta6220 PMID:12433315

Klier, C. M., Muzik, M., Rosenblum, K. L., & Lenz, G. (2001). Interpersonal Psychotherapy adapted for group setting in the treatment of postnatal depression. *The Journal of Psychotherapy Practice and Research*, *10*, 124–131. PMID:11264336

Marks, I. M., Cavanagh, K., & Gega, L. (2007). *Hands-on Help: Computer-aided Psychotherapy*. Psychology Press Ltd.

Marks, I. M., Shaw, S. C., & Parkin, R. (1998). Computer-aided treatments of mental health problems. *Clinical Psychology: Science and Practice*, *5*(2), 151–170. doi:10.1111/j.1468-2850.1998.tb00141.x

McCarthy, M., & McMahon, C. (2008). Acceptance and Experience of Treatment for PND in a Community Mental Health Setting. *Health Care for Women International*, *29*, 618–637. doi:10.1080/07399330802089172 PMID:18569047

Merry, S. N., Stasiak, K., Shepheard, M., Frampton, C., Fleming, T., & Lucassen, M. F. G. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *BMJ (Clinical Research Ed.)*, *344*, 1–16. PMID:22517917

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Milgrom, J., Ericksen, J., McCarthy, R., & Gemmill, A. W. (2006). Stressful impact of depression on early mother–infant relations. *Stress and Health*, *22*(4), 229–238. doi:10.1002mi.1101

Morrell, J. C. (2006). Review of interventions to prevent or treat PND. *Clinical Effectiveness in Nursing*, 9(Suppl. 2), e135–e161. doi:10.1016/j.cein.2006.11.006

NICE. (2006). Computerised cognitive behavioural therapy for depression and anxiety. Retrieved December. 2015. Technology appraisal guidance. http://www.nice.org.uk/nicemedia/pdf/ccbtassessmentreport.pdf

O'Mahen, H. A., Richards, D. A., Woodford, J., & Wilkinson, E., McGinley. J., Taylor, R. S., Warren. (2013). Netmum, a phase II randomized controlled trial of guide Internet behavioural activation treatment for postpartum depression. *Psychological Medicine*. PMID:24148703

Pampallona, S., Bollini, P., Tibaldi, G., Kupelnick, B., & Munizza, C. (2002). Patient adherence in the treatment of depression. *The British Journal of Psychiatry*, *180*(2), 104–109. doi:10.1192/bjp.180.2.104 PMID:11823317

Payne, K. A., & Myhr, G. (2010). Increasing access to cognitive–behavioral therapy (CBT) for the treatment of mental illness in Canada: A research framework and call for action. *Healthcare Policy*, *5*, e173–e175. PMID:21286263

Proudfoot, J., Ryen, C., Everitt, B., Sharpio, D., Goldberg, D., Mnn, A., ... Gray, J. A. (2004). Clinical efficacy of computerized cognitive-behavioural therapy for anxiety and depression in primary care: Randomised controlled trial. *The British Journal of Psychiatry*, *185*(1), 46–54. doi:10.1192/bjp.185.1.46 PMID:15231555

Rojas, G., Fritsch, R., Solis, J., Jadresic, E., Castillo, C., González, M., ... Araya, R. (2007). Treatment of postnatal depression in low-income mothers in primary-care clinics in Santiago, Chile: A randomized controlled trial. *Lancet*, *370*(9599), 1629–1637. doi:10.1016/S0140-6736(07)61685-7 PMID:17993363

Ruaro, L. (2013). *The experience of mother who received therapy of PND. A qualitative study* [Professional Doctorate in Counselling Psychology].

Turner, K. M., Chew-Graham, C., Folkes, L., & Sharp, D. (2010). Women's experiences of health visitors delivered listening visits as a treatment for postnatal depression: A qualitative study. *Patient Education and Counseling*, *78*(2), 234–239. doi:10.1016/j.pec.2009.05.022 PMID:19574015

Wan', J., Hu, B., Moore, P., & Ashford, R. (2008). Intelligent mobile computing to assist in the treatment of depression. *Pervasive Computing and Application, IEEE*, *2*, 650–655.

World Health Organisation. (2003). *Adherence to long-term therapies: evidence for action*.

Yonkers, K. A., Lin, H., Howell, H. B., Heath, A. C., & Cohen, L. S. (2008). Pharmacological Treatment of Postpartum Women with New Onset Major Depressive Disorder: A Randomized Controlled Trial with Paroxetine. *The Journal of Clinical Psychiatry*, *69*(4), 659–665. doi:10.4088/JCP.v69n0420 PMID:18363420

This research was previously published in the International Journal of Public Health Management and Ethics (IJPHME), 2(1); edited by Qiang (Shawn) Cheng; pages 1-18; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Questionnaire

1. What therapy did you use for postnatal depression? (*Please tick all that apply*)
    - Cognitive Behavioural Therapy
    - Monitoring and assistance (by a GP, Nurse or Therapist)
    - Medication
    - Psycho-education and information
    - Self-help from peer group
    - Interpersonal Therapy
    - Well-being Therapy
    - Problem-solving Therapy
    - No treatment
    - Other
2. Which ONE of these therapies did you find MOST effective for postnatal depression? (Please choose one)
3. How effective was the therapy you chose in question 2 in addressing the following?
    - It helped to reduce my isolation
    - It helped me talk to other people
    - It helped me prioritise demands on things I have to do
    - It helped my relationship with my children
    - It helped reduce my stress
    - It helped me to manage negative thoughts
    - It helped me to have faith in myself
4. How long did it feel you had to wait between when you were diagnosed with postnatal depression and when you started receiving therapy? (Please choose one)
5. When you had appointments to receive therapy, how convenient was the location of the facility for you?
6. How affordable was the cost of therapy to you? Costs include bus fare, travel costs, child care or time of work.
7. Thinking about the time you spent on each appointment, how acceptable did you find the length of time spent receiving treatment
8. How long did you have to use your most successful therapy for postnatal depression until recovery?
9. How often did you follow instructions specific to using your therapy when you had postnatal depression?
10. How frequently did you use your most successful therapy for postnatal depression?
11. How long did you have to use your most successful therapy for postnatal depression until recovery?
12. How often did you follow instructions specific to using your therapy when you had postnatal depression?

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 48
Modelling and Simulating Complex Systems in Biology:
Introducing NetBioDyn – A Pedagogical and Intuitive Agent-Based Software

Pascal Ballet
Université de Bretagne Occidentale, France

Karine Pichavant
Université de Bretagne Occidentale, France

Jérémy Rivière
Université de Bretagne Occidentale, France

Frank Abautret
Collège Max Jacob, France

Alain Pothet
Académie de Créteil, France

Alexandra Fronville
Université de Bretagne Occidentale, France

Michaël Theron
Université de Bretagne Occidentale, France

Vincent Rodin
Université de Bretagne Occidentale, France

ABSTRACT

Modelling and teaching complex biological systems is a difficult process. Multi-Agent Based Simulations (MABS) have proved to be an appropriate approach both in research and education when dealing with such systems including emergent, self-organizing phenomena. This chapter presents NetBioDyn, an original software aimed at biologists (students, teachers, researchers) to easily build and simulate complex biological mechanisms observed in multicellular and molecular systems. Thanks to its specific graphical user interface guided by the multi-agent paradigm, this software does not need any prerequisite in computer programming. It thus allows users to create in a simple way bottom-up models where unexpected behaviours can emerge from many interacting entities. This multi-platform software has been used in middle schools, high schools and universities since 2010. A qualitative survey is also presented, showing its ability to adapt to a wide and heterogeneous audience. The Java executable and the source code are available online at http://virtulab.univ-brest.fr.

DOI: 10.4018/978-1-7998-1204-3.ch048

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The theory of complex systems has become more and more relevant over the last thirty years to understand concepts such as emergence and self-organization and their role in biological, physical, chemical and social systems (Jacobson & Wilensky, 2006). Modelling and simulating these concepts behind the theory of complex systems is now very important, especially for biologists (students, teachers, researchers). For example, recent works have highlighted the complementarity of modelling both bottom-up and top-down (aggregated) approaches (Stroup & Wilensky, 2014). It has been shown though that teaching and learning complex systems are challenging tasks, and raise a lot of pedagogical issues. The two main issues are:

1. A widely taught top-down approach that have led students to think in that unique way; and
2. System dynamics that are difficult to understand and can be non-linear and counter-intuitive (Jacobson & Wilensky, 2006).

Among different strategies, the use of Multi-Agent Based Simulations (MABS) in research, teaching and learning has proved to be an efficient way of answering (some of) these issues (Epstein, 1999; Jacobson & Wilensky, 2006; Ginovart, 2014). Indeed, the individual-based approach, in opposition to a population-based approach (such as ordinary differential equations), focuses on the entities of the system, their behaviour and their local interactions to explain the global system's behaviour. As such, this concept is easier to understand, more intuitive and do not require advanced mathematical skills (Ginovart, 2014). The use of computer tools to implement these models in virtual environments and simulate them allows researchers and students to generate hypothesis, test them and, especially for students, build their knowledge more efficiently (in a process of *abduction* (Houser, 1992)), see Section 2.2). NetLogo (Wilensky, 1999) is a good example of such agent-based software, and has been used in a lot of works, both in education (Gammack, 2015) and research (Banitz, Gras, & Ginovart, 2015).

This chapter focuses on how to teach complex biological systems to, for example, university students in Biology, middle school and high school students; students who often do not have any experience in computer programming, and/or are reluctant to implement any code. The problem is that most of agent-based software require knowledge and skills in programming languages such as Java (*e.g.* the Repast software (North, Collier, & Vos, 2006)) or the NetLogo environment. This is an issue that occurs more and more as the use of agent-based software, both in research and teaching, grows in social and life sciences. In (Ginovart, 2014) and (Gkiolmas, Karamanos, Chalkidis, Skordoulis, & Papaconstantinou, 2013), a model of a predator-prey system is used with NetLogo to help respectively first-year and high-school students to understand individual-based approaches and eco-systemic mechanisms. Despite its interest, this model is a ready-made one, coded in NetLogo, where the students can change some parameters of the simulation, but are not implied in the process of creating their own simulation from scratch. Indeed, implementing this model is considered as "a hard task" by the students (Ginovart, 2014) and could be counterproductive: looking for errors in the code instead of thinking at the rightness of the model. The authors argue that using pre-existent models could have a part in preventing students for building their own system's *representation* (this point will be developed in Section 2.2). The gap between most of agent-based software prerequisites and the actual programming skills of students (and sometimes researchers) in these fields has to be dramatically reduced and the processes of implementing a model and simulate it should be more intuitive.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This chapter essentially looks at the modelling of biological systems at and around the cellular scale. However, it is interesting to note that a lot of works have been focusing on modelling and simulating systems at the molecular scale, resulting in several software developed for students. For example, it is possible to find molecular visualisation systems like PyMol (Schrödinger, 2015) or software that help the docking of molecules like AutoDock Vina (Trott & Olson, 2010). There is also an automatic plotting of protein-ligand interactions with Lig-Plot (Wallace, Laskowski, & Thornton, 1996). It is possible to visualise molecular interaction networks and biological pathways thanks to software like CytoScape (Shannon et al., 2003). When focusing on DNA, a complete suite in Python can be used for the sequence alignment (BioPython - Cock et al., 2009). All these software are sources of inspiration when dealing at the cellular level.

In this chapter, the authors propose a software called NetBioDyn working like a game construction kit, where entities are located in a 3D grid environment and follow simple local transformation rules. In the name NetBioDyn, Bio stands for Biology, Dyn for Dynamic and Net comes from Internet because initially NetBioDyn was a Java Applet. Though this software is dedicated to cellular systems where several cells are in interaction, it can also model and simulate molecular mechanisms in a simple and symbolic way. The process of building models and simulations is facilitated by a graphical user interface similar to a drawing software like Microsoft Paint. Researchers or students can build their own models, continuously visualize a clear representation of these models, and see them evolving in time and space. The results are displayed like a chess board with moving and changing coloured entities and can be analysed thanks to recorded population curves. From the point of view of the researcher (who usually is a teacher too), the software can help to explain non-intuitive and complex biological systems in a visual way, easily understandable by his peers or students. The authors present in this chapter the key principles of the software in rigorous but comprehensive language and develop three examples coming from real biological systems. The first example is a predator-prey system involving two bacteria. The second one simulates the mechanism of oxygen exchange in the zebra fish gills between water and blood. The third one simulates the blood coagulation mechanisms in a small section of a vein. A discussion is then engaged to show the interests of such software as well as the limits and drawbacks. A qualitative survey is also presented to show how students in biology and students in computer sciences have received the software. After the conclusion, different perspectives are presented in order to improve both the software and its use.

BACKGROUND

The use of computers in teaching biology can be divided into three common practices:

1. The development of models and simulations using a programming language like C++, Python or Java. This practice is powerful but takes time and requires advanced programming skills.
2. The use of customizable software, where students work on pre-existent models and can only change some parameters in simulations. This approach is very easy to use with students, but is not as efficient as building their own models.
3. The use of modelling software to help students develop their own models without any programming skills.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This third practice requires though the development of complex software that simplifies the process of modelling, by proposing for example, a graphical representation and construction of virtual agents, their behaviours and their environment.

Naturally, these three practices can also be found in the use of Multi-Agent Based Simulations and agent-based software in teaching biology. A quick overview of:

1. Agent-based modelling software that allow the development of models and simulations thanks to a programming language,
2. Agent-based modelling software that propose pre-made simulations, or
3. Computer aided-design modelling software is presented here.

A more complete review of agent-based environments has been made by Nikolai and Madey (2009), where more than 50 platforms are classified and discussed. Readers who wish to deepen the potential of the multi-agent approach for the simulation of complex systems can refer to the book edited by Weiss (2000), especially the chapter "Search Algorithms for Agents".

Agent-Based Software in Biology

Programming Languages

NetLogo (Wilensky, 1999) is one of the most well-known software allowing the development of agent-based programs, including simulations. It has a simplified programming language and simple mechanisms to create interfaces. This software is widely used in the community of biologists and many biological systems have been developed thanks to it. It has also been specifically used to help students understand the mechanisms behind complex systems (Stroup, 2014, Jacobson & Wilensky 2006). Nevertheless, pupils or students have to learn a programming language if they want to create their own models (Ginovart, 2014, Gkiolmas et al., 2013). One can also mention AgentSheets (Repenning, & Sumner, 1995) and AgentCubes (Repenning, Smith, Owen, & Repenning, 2012), which are simple programming environments allowing middle school pupils to learn programming in an appealing manner. These software are designed for multi-purpose applications and mostly games. They are based on a multi-agent system where behaviours are created by drag and drop condition-action algorithms. Although the programming is easy, some code is still required. Among generic multi-agent platforms, one can cite Breve (Klein, 2014; Klein & Spector, 2009) and Repast (North et al., 2006). Despite the fact that they have integrated development environments, the creation of models and simulations needs the knowledge of a programming language like Python, Groovy or Java. More recently, Biocellion (Kang, Kahan, McDermott, Flann, & Shmulevich, 2014) allows accelerating discrete agent-based simulation of biological systems with millions to trillions of cells. It uses cluster computers by partitioning the work to run on multiple compute nodes. It requires though mathematical modelling backgrounds and C++ programming skills.

Pre-Maid Simulations

Different simulations of multi-agent systems have been developed with NetLogo. The community models website of NetLogo proposes, for example, a customizable simulation of a predator - prey - poison system with 11 parameters: initial number rabbits and coyotes, rabbit reproducing probability, grass regrowth

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

time, poison quantity, etc. (Jenkins, 2015). Concerning the simulation of multicellular systems, one can cite VirtualLeaf (Merks, Guravage, Inzé, & Beemstaer, 2001) which allows the simulation of tree-like structures based on agent representing cells and regulation network. It includes 12 different models and about 60 parameters that can be adjusted (e.g. cell mechanic or mathematical integration parameters).

Computer Aided-Design Software

In addition to a simplified programming language, NetLogo provides a "System Dynamic Modeler" that can be used to represent the dynamic of the system, entities inter-relations, feedback loops, flows and stocks. However, this tool requires a global understanding of the system (that is often missing with an agent-based approach) and is quite difficult to use as an intuitive modelling tool. Ioda (Interaction-Oriented Design of Agent simulations) and Jedi (an implementation of the Ioda concepts), allow the creation of multi-agent simulations in a very simple manner (without any coding skill), where predefined interactions can be added between agents (Kubera, Mathieu, & Picault, 2008). As Jedi is a demonstration software, it cannot be used to create complex simulations (4 agents maximum can be manipulated at the same time and there is no graphical user interface to precisely build the initial state). AnyLogic (XJ Technologies Company Ltd., 2012) is a multipurpose simulation software with a visual development environment. It has pre-built object libraries allowing users to create models by a drag-and-drop of elements. Its generic architecture makes it a powerful software but it could be relatively complex to understand for some students, especially the secondary ones. Another very interesting human-computer interaction is developed in the SimSketch software (Bollen, 2013). In this approach, the user draws the agents using an integrated paint tool then adds a name, a type and different pre-defined behaviours. This software focuses on primary and early secondary school students whereas this work focuses on the end of the secondary school to the university. Note that, this tool is not yet totally usable to create multicellular systems because there is no if...then...else structure. The software called CompuCell3D (Izaguirre et al., 2004) allows the simulation of multicellular systems with a user-friendly interface. The Python programming language can be used to create models but an XML file can also be edited to create simulations without coding. SimBioDyn (Ballet, Zemirline, & Marcé, 2004) is a software based on a hybrid multi-agent and mass-spring system designed to graphically create multicellular systems. Although rich in features (cell deformation, migration, adhesion and differentiation, molecular production and consumption and configurable through scripts in Python), it remains hard to use by non-specialists without a course of several hours (Ballet et al., 2004).

Among computer aided-design software, there are also some interesting works that have focused on modelling biological systems following other bottom-up or top-down approaches. For example, one can cite Virtual Cell (Shin, Liu, Loew, & Schaff, 1998), a software allowing to model intra-cellular mechanisms thanks to differential equations. It has a rich web-based Java interface to specify compartmental topology and geometry, molecular data and interaction parameters. However, the purpose of this software is to simulate a single whole cell or sub-parts of a cell but not multicellular systems, and it is not fitted for non-specialist users as advanced skills in mathematics are needed.

In this chapter, the authors focus on teaching pupils or students with no programming knowledge complex biological systems, such as multicellular systems.

In this context, the interests of using computer aided-design software are multiple:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- The student do not focus on how to program but on what to model (they remain in the field of biology).
- They cannot make syntactic or lexical errors: their models are always functional despite the fact that they can make semantic errors.
- They are no longer in charge of just some simulation parameters, but become the main actors of the model building, thus favouring the building of their knowledge.
- The teacher visualizes quickly where a student has insufficient knowledge.

So what is needed is a modelling software designed to be used by students with no programming knowledge and dedicated to agent-based multicellular systems modelling and simulation. However, this survey shows that many computer aided-design software exist, but to our knowledge none of them are intuitive enough to be used in this context. In order to answer these issues and give students the ability to easily create their own models, the authors have developed the software called NetBioDyn.

Reasoning About Complexity

Presenting essential concepts on system dynamics has become very important for students in biology. Teaching and reasoning about complex systems still raise a lot of pedagogical issues as shown in particular by Jacobson and Wilensky (2006).

Lynch and Ferguson (2014) argue that grasping the complexity requires a representational framework able both to describe executable computer models *and* to support reasoning as a simply readable "external representation", as opposed to internal representations that are the user's mental models (Zhang and Patel, 2006). According to them, a good external representation affords the classic scientific approach, called *abduction*, emphasizing the observation, the build of hypothesis and the tests to verify the relevance of assumptions (Houser, 1992): it promotes the exploration and understanding of a phenomena thanks to a practical approach. It should be too constructed by the students themselves for two principal reasons: the first one is that it improves their learning (better memorization, quasi-physical experimentation of the model and appropriation of fundamental mechanisms) and the second one concerns the amelioration of their self-commitment (Prain, & Waldrip, 2006). Lynch and Ferguson define an ideal representation as "some written, external representation to complement user's internal representations [...], producing an external cognitive artefact which facilitates exploration and discovery [...] while also offering the ability to be executed".

Lynch and Ferguson summarize the 9 properties of a suitable external representation that were previously defined by Zhang and Patel (2006):

- Short-term or long-term memory aids.
- Directly perceived information.
- Knowledge and skills that are unavailable internally.
- Support for easy recognition and inference.
- Unconscious support for cognitive behaviour.
- More efficient action sequences generated.
- Facilities to stop time and make information visible and sustainable.
- Need for abstraction reduced.
- Aids to decision making through high accuracy and low effort.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Following these properties, the authors claim that NetBioDyn, thanks to its graphic interface and its simulation tools proposes a suitable external representation. It gives to students and teacher a simple way to create and experiment and test models of biological systems, while providing at any time a simplified and complete view of the system's state.

More generally, readers who want to improve their knowledge on "learning by doing" in virtual environments (VE) can read the book of Aldrich (2009).

NETBIODYN, A PEDAGOGICAL AND INTUITIVE AGENT-BASED SOFTWARE

NetBioDyn is a fully open source Integrated Development Environment (IDE) dedicated to teaching activities in biological complex systems, mainly at the cellular level and secondarily at the molecular scale. It gives the possibility to teachers and students to create simulations of multicellular and multi-molecular systems in a simple and visual manner. NetBioDyn does not require any skill in computing and lays on the agent-based approach. Besides being simple to understand for students, this approach allowed the authors to design an intuitive graphical user interface to easily manipulate the agent-based concepts: entity, behaviour and interaction.

In this section the graphical user interface of the software is first presented. Then, the section goes inside the software architecture with the description of the entities populating the simulated environment and how they can evolve thanks to algorithmic behaviours (movement, differentiation, division, apoptosis, *etc.*). In the next section, three examples of easily reproducible simulations are detailed. The code and the executable (for UNIX, Mac and Windows Operating Systems) are available for download on the website http://virtulab.univ-brest.fr.

Graphical User Interface

For educational purpose, the design of a graphical user interface is a crucial point. For the student or the teacher who creates a simulation it must be:

- Intuitive, to prevent the user from being lost in multiple buttons or lists.
- Logic, to build a structured simulation.
- Clear, to show important data and hide less important ones.
- Accessible, to concentrate on "what to simulate" instead of "how to simulate".
- Complete, to have a powerful tool that can simulate advanced biological behaviours.
- Attractive, to stimulate student's imagination and motivation.

For the teacher who uses the software in a course, it must:

- Be robust, avoiding the teacher to help students in case of inappropriate action (for example, no source code is required, avoiding student to make syntax mistakes - only semantic errors can be made, which are the most relevant ones to test the student comprehension of the modelled biological system).
- Show relevant information, to quickly see where each student is and if he is mistaken.
- Allow a short learning curve, not to spend more than one hour to explain how the software works.

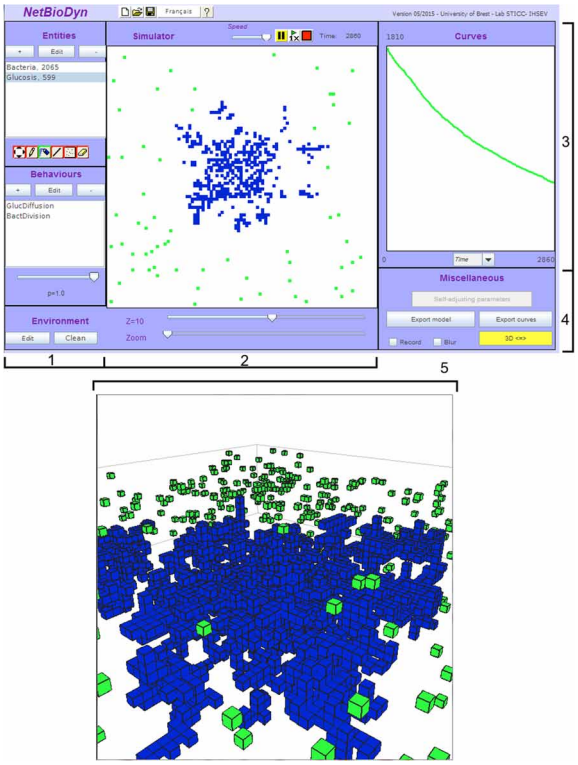
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

To try to achieve these points, the authors decided to use Multi-Agent Based Simulations (MABS) to guide the development of the graphical user interface. MABS are naturally close to how biological objects like cells are described, reducing the abstraction needed to make simulations. In this chapter, the authors use the term "entity" to represent biological objects (or group of objects) and the term "behaviour" to explain how the entities evolve in time. The word "environment" is used to represent the space where simulations occur. A good analogy can be done with board games like chess: an entity is a piece, a behaviour is a game rule, the environment is the board and the initial state of the simulation corresponds to the initial positions of the pieces.

Figure 1 shows the different parts of the interface. The first part of the graphical user interface is made of a list of entities, followed by a list of behaviours and then a button to change the environment's properties (such as size and textual description). The position order of the two lists is logic: entities must be defined before their behaviours. Nevertheless, note that a new entity or a new behaviour can be added, changed, renamed or removed at any time by the user to support the creation process which is rarely linear. The environment button is located below the two lists because the default properties do not change frequently.

The second part of the graphical user interface is the placement and the display of entities in their environment. A central panel serves as a paint area where entities can be individually or collectively located or removed thanks to classic drawing tools like pen, spray, line and eraser, plus a random initializer tool. Entities are located in several layers that can be displayed or hidden using the Z slider. The simulator is controlled by three buttons:

Figure 1. NetBioDyn graphical user interface



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Play (a green right arrow),
- Play one step (an orange right arrow), and
- Stop (a red square).

The simulator can also be paused at any time to observe a specific state or make a screen capture, and its speed can be reduced or increased.

The third part displays the population curves of entities evolving in real time. Several curves can be displayed at the same time to compare their values according to the simulation steps. The user can also display several population curves according to a specific entity (for an example, see next section Examples).

The fourth part enables the user to change miscellaneous properties and export simulation data in the well-used formats in Biology CSV (spreadsheet) and R programming language. A tool to automatically adjust the simulation's parameters according to *in-vitro* results is a work in progress and will be available soon.

The fifth and last part display in a specific frame the environment in 3D in real-time.

In the next subsection, the NetBioDyn entities are described in details.

Entities

An entity in NetBioDyn is a simple cube of size 1 x 1 x 1, located as a grid element of the environment (which is described in details below), having a unique name and a unique colour. Other forms (triangle, disc, star) or images can also replace the coloured cube.

An entity also has a half-life parameter indicating its speed of degradation or destruction.

For each entity, a short description can also be added. This is made for two main reasons. The first is to encourage students to comment their models because they will have to maintain and share their work. The second reason is to help students to understand pre-made models that the teacher created.

Finally a cleanable property can be unselected to prevent entities from being removed when the environment is cleaned by the user. This is useful when the teacher created a simulation where the initial state contains a minimum of entities located at specific positions: the student can clean his own entities but the teacher's ones always remain.

The entity configuration window can be seen on Figure 2.

In the following sub-section the behaviours allowing entities to change their states are explained.

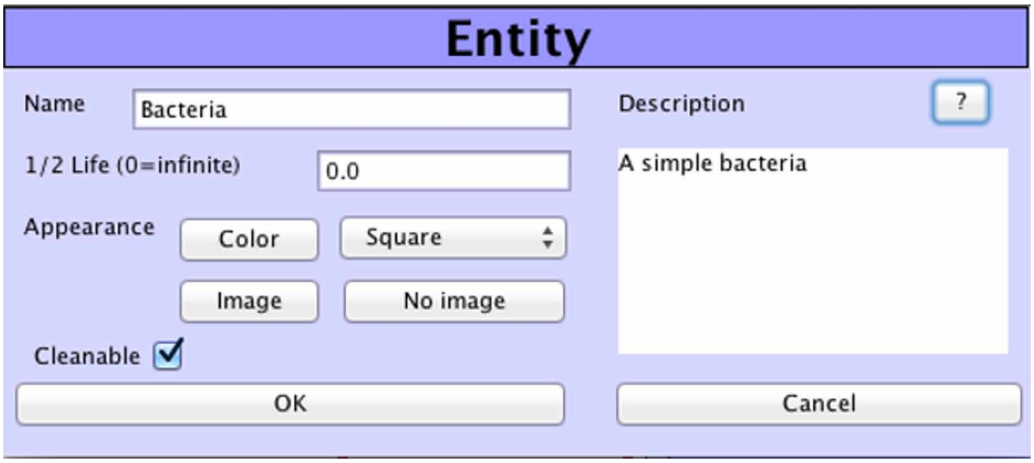
Behaviours

In NetBioDyn, a behaviour aims to change entity's states. For example an entity can be put in motion, differentiated or destroyed. Four parts are required to define a behaviour, three conditions and one action:

- Conditions
  - Which entities are involved?
  - Where are they located from each other?
  - Is the probability validated?
- Action
  - What will happen to them?

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The name, the half-life and the appearance of an entity is made in the entity configuration window. Note also the description field on the right and the cleanable check-box at the bottom-left



The first condition identifies all entities that are involved in the simulation. Only those that are in a good position from each other are kept (condition two). Thanks to the third condition, a behaviour is probabilistically performed: a random number between 0.0 and 1.0 is generated and must be inferior to the probability of the behaviour.

If the conditions are validated at different areas in the environment, a behaviour applies several times during one simulation step. Conflicting areas are detected where a same entity can be used two or more times by behaviours. In this case, only one area is randomly chosen among the conflicting areas (see Figure 3).

In order to ease the design of behaviours, the authors created a graphical representation consisting of two main parts (see Figure 4). The first part containing three lines indicates the behaviour's name, its description and its probability to occur. The second part is divided in three columns: the left one contains the entities involved in the behaviour (reactive), the middle one indicates where the entities are located one from each other and the right column indicates how entities are changed.

In the software NetBioDyn, the graphical user interface is very close in terms of displayed data. The entities are selected thanks to a combo-box containing all the created entities. The positions are selected by clicking on the light or dark squares. The cross at the left allows the user to remove a full line of reactive and products (see Figure 5).

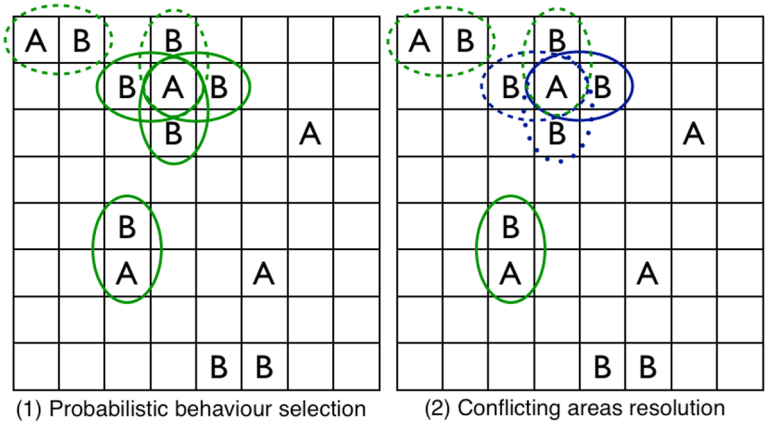
The next subsection describes the environment, where entities are placed and behaviours occur.

Environment

The environment represents a 2D or 3D space where entities evolve and behaviours take place. It is a simple grid where a grid cell contains none or a maximum of one entity. An entity occupies a whole grid cell. It follows Von Neumann neighbourhood: it has four neighbours in 2D and six in 3D. The size of the environment is by default 100 x 100 x 2 but can be changed by the user. The view is only 2D when the user put, move or remove entities but a 3D view can be activated for a better and more intuitive rendering.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. In (1), a behaviour can be applied several times during one simulation step. In this example the system detects that the entities A and B are in contact. The probability to perform the behaviour, if the entities are found in contact, is 0.8. So, of the six areas where the behaviour can be performed, four are kept (continuous circles) whereas two are not (dotted circle). Note that this area configuration can change accordingly to the random numbers generated. Moreover notice that three areas are in conflict and must be treated before the action can be applied (continuous circles including the same entity A). In (2), the system manage the conflicting areas. Among them, only one can be kept (the horizontal continuous circle) whereas the two others cannot. All conflicting areas have the same probability to be kept (or not kept). Note that the probabilities of behaviours have already been validated before the conflicting areas are treated. Finally, only two areas will apply the behaviour (continuous circles)



From the principles presented, the next section develops three examples that can be realised by pupils or students.

EXAMPLES

The examples detailed in this section can be performed using the software NetBioDyn freely available at http://virtulab.univ-brest.fr.

It can be made by pupils from age 14 to university students. The first example deals with a prey-predator system involving two species of bacteria: Bdellovibrio and Photobacterium leiognathi. The second example models the dioxygen exchanges between water and the blood of zebra fishes. Finally, the third one aims to reproduce and understand the blood coagulation mechanisms.

These examples show to the students how to create a model in four stages:

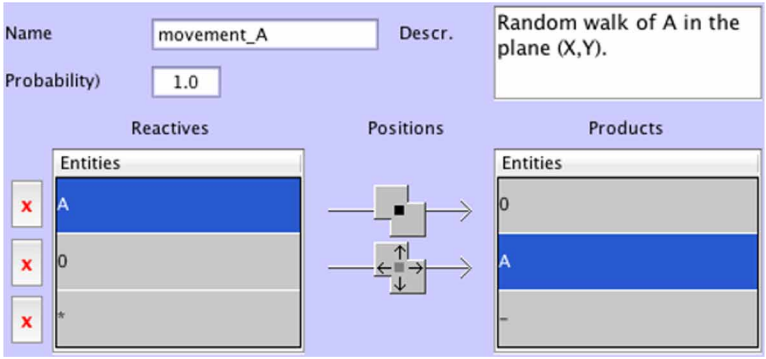
- **The Real System**
  - What is the current knowledge?
  - What part of the system would I model?
  - What are the limits of my study (dimensions, time, energy, *etc.*)?
  - Am I able to distinguish, among the current knowledge, what are hypothesis and what are facts?

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. For this example, the behaviour is read like this: "the entity A, which is the centre of the behaviour (+) must be in contact with one entity B and one entity C. B can be anywhere (up, down, left, right, front, behind) around A and C can be at the left, the front or behind A. Moreover, if the probability is validated and the possible conflicts resolved, A will become A', B will become B' and C will be changed in C' ". The first reactive is always at the centre of the reaction. A reactive or a product can be 0 which means no entity. The diagonals are not taken into account in the positions

Name	**behaviour**	
Description	Short text of explanation	
Probability	Decimal number in [0, 1]	
Reactives	Positions	Products
A	+	A'
B	⊙ ↑ ← A → ↓ ⊗	B'
C	↑ ← A ↓	C'

*Figure 5. The graphical user interface of behaviours in NetBioDyn has the same properties as seen previously. In this example, an entity A is removed from its current position (first line of the lower part), to be placed at a free location (second line). The third line contains the symbol * as reactive (which means anything) and the symbol - as product (which means do nothing). This behaviour applies a random walk to all A entities in the simulation*



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **The Simplifications**
  - What have I to keep absolutely from the real system?
  - What can be neglected?
- **The Model**
  - How to represent real entities using virtual ones?
  - How to model biological mechanisms thanks to algorithms or equations?
  - What are the dimensions and the topology of the environment and how to represent it on a computer?
- **The Simulation**
  - Can I trust the results?
  - What are the limits concerning explanation and prediction?
  - Are the results qualitative or quantitative?
  - Does the simulation inspire me in order to prepare real experiments?
  - Is the simulation robust to small changes (probabilities or initial state for example)?
  - Is the simulation clear enough to share and confront my knowledge with others?

The next sub-section expose in details the process of modelling three cellular and molecular systems, following these four stages from the real system to the simulation.

Bdellovibrio and Photobacterium Leiognathi

Real System

Bdellovibrio and Photobacterium leiognathi are marine bacteria. The interactions between these two bacteria represent a good example of a predator-prey system. Bdellovibrio is a predator of Photobacterium leiognathi. It penetrates inside its prey, proliferates, kills the host and finally releases its progeny. This cycle time is about 4 hours. Bdellovibrio is a bacterium of 1 μm in size and thanks to its flagellum, moves very fast (up to 160 μm/s). The size of Photobacterium leiognathi is about 2 μm. They usually are in symbiosis with fishes and, in this case, have not shown to be motile.

Biological systems involve too many parameters, objects and behaviours to be fully modelled and simulated. A simplification of the system described here is needed in order to focus on the predator-prey principles.

Simplification

In order to reproduce the Bdellovibrio and Photobacterium leiognathi relationship, a simplified model which takes into account five objects and five behaviours is proposed:

- Biological entities:
  - Bdellovibrio (Bdel).
  - Photobacterium leiognathi (Leio).
  - Photobacterium leiognathi infected (LeioInfected).
  - Fish tissue where Photobacterium leiognathi lives (FishT).
  - Nutriment for Photobacterium leiognathi coming from the fish (Nutri).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Biological behaviours:
  - Bdellovibrio movement (BdelMvt).
  - Infection of Photobacterium leiognathi by Bdellovibrio (LeioInfection).
  - Photobacterium leiognathi lyse and Bdellovibrio release (LeioLyse).
  - Photobacterium leiognathi division (LeioDiv).
  - Nutriment production for Photobacterium leiognathi (NutriProd).

This model has to be also adapted to our simulator. The next section explains how to put the model in the software NetBioDyn, which has its own possibilities (*i.e.* multiple entities and behaviours) and constraints (*i.e*. the environment is a 3D grid and a grid element has a maximum of six neighbours).

Model

The entities chosen for the simulation are Bdel, Leio and LeioInfected. The fish tissue and the nutriments are implicit in our simulation. This is a choice which can be changed if the fish tissue or the nutriments must be quantitatively taken into account or if their spatial distribution has to be studied. The graphical representation used for entities are simple coloured cubes of volume 1 μm^3. In this example Bdel is red, Leio is blue and LeioInfected is purple. Bdel and Leio have a half-life of 100 simulation steps (which means that after 100 simulation steps a Bdel or a Leio entity has 50% of chance to die, or said differently, after 100 simulation steps for a non-reproductive population of N individuals, only N/2 individuals will remain).

Four behaviours have been kept to simulate the bacterial system (see Figure 6):

1. The movement of Bdel entities which is a simple 3D random walk. The probability of this behaviour is 1.0, the maximum possible, because no other behaviour is as fast as this one. The bibliography gives the speed of 160 μm/s for Bdellovibrio. This is a maximum and an instantaneous speed. Due to its frequent directional changes, its average speed is closer to few μm/s. A grid cell can contain one bacterial entity at the same time, so a grid cell represents a cube of 1 x 1 x 1 μm^3.
2. The infection of Leio by Bdel. The probability of infection is 0.1, indicating that one contact over ten can produce an infection.
3. The lysis of Leio and the release of two Bdel. The probability is chosen to reproduce the 4 hours of incubation for Bdel. A time step is equivalent to one second, so 4 hours are simulated by 14400 steps. At this time, this is supposed that 99% of the LeioInfected have been killed.
4. The division of Leio following the nutriment intake.

The environment is a 100 x 100 x 2 grid. All the Leio entities are at z=0 in order to imitate the fact that they are on the fish tissue.

Simulation and Results

The initial state of the simulation is given by Figure 7. The initial state is made of 1000 Leio entities and 100 Bdel entities randomly located in the environment. Leio entities have been placed at z=0 while Bdel entities are located at z=1. The number of preys is ten times inferior to the number of predators in order to initiate the system without having the preys eaten by too numerous predators at the beginning. They all are located at random in the same 3D environment.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The four behaviours described using the previously seen representation

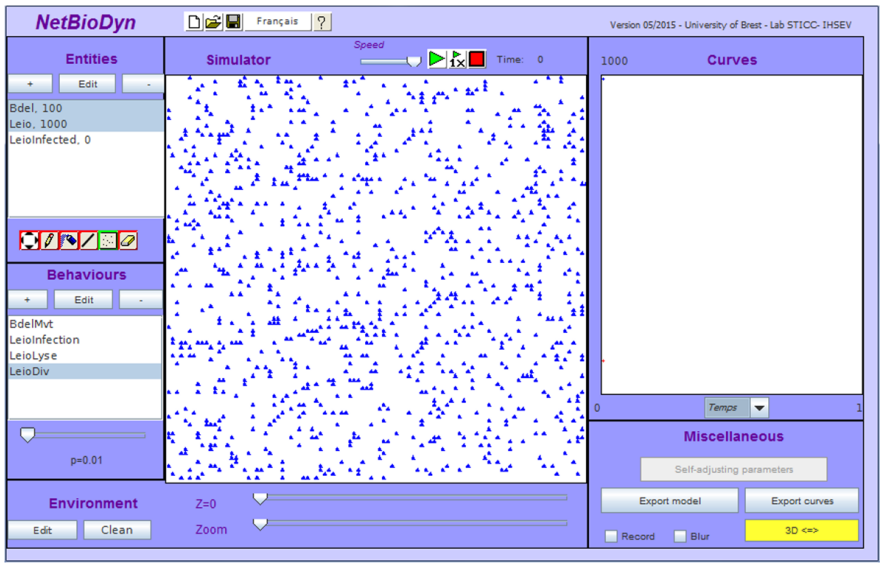
Name	BdelMvt	
Description	Random walk of Bdellovibrio bacteria.	
Probability	1.0	
Reactives	Positions	Products
Bdel	+	0
0	⊙ ↑ ← Bdel → ↓ ⊗	Bdel
*		-

Name	LeioInfection	
Description	Infection of photobacterium Leiognathi by Bdellovibrio.	
Probability	0.1	
Reactives	Positions	Products
Leio	+	LeioInfected
Bdel	⊙ ↑ ← Leio → ↓ ⊗	0
*		-

Name	LeioLyse	
Description	Lyse of photobacterium Leiognathi and release of two Bdellovibrio.	
Probability	0.1	
Reactives	Positions	Products
LeioInfected	+	Bdel
0	⊙ ↑ ← LeioInfected → ↓ ⊗	Bdel
*		-

Name	LeioDiv	
Description	Division of photobacterium Leiognathi following the nutriment intake.	
Probability	0.01	
Reactives	Positions	Products
Leio	+	Leio
0	↑ ← Leio → ↓	Leio
*		-

Figure 7. This screen capture of NetBioDyn shows the initial state of the simulation where 1000 Leio (triangles) and 100 Bdel (squares) have been placed at random in the environment using the random initializer tool. Only entities placed at z=0 are here visible

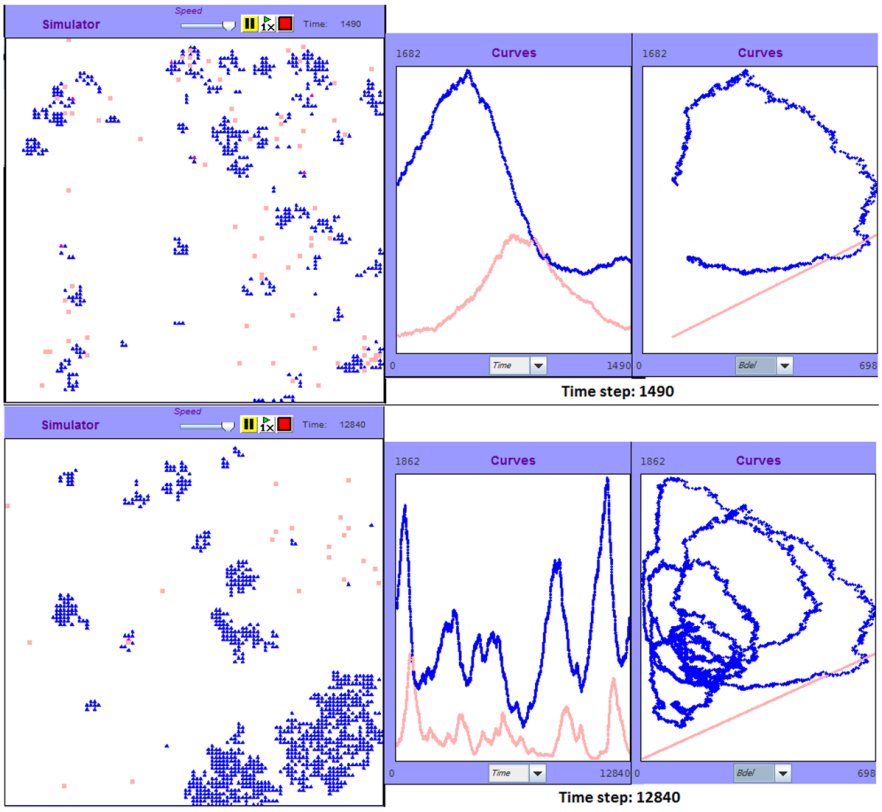


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8 shows two snapshots of the simulation at time steps 1490 and 12840. At time step 1490 (upper screen capture), the first predator-prey cycle is shown, where the Leio population (preys) grows initially, just followed by the growth of the Bdel population (predators). Then, the Leio population decreases because they are eaten by the Bdel, and then the Bdel population decreases too by lack of food, resulting in an increase of Leio due to the fact there are less Bdel entities. The two peaks of the curves are shifted in time: the peak of the preys (Leio) occurs before the peak of predators (Bdel) because the predators need the preys to grow. Small spots of Leio entities appear because of their bacterial division. Five other cycles are shown at the lower part of the figure (time step 12840). The stochastic aspect of the simulation is well shown in the irregularities of curves and entity positions. The predators are close to the preys not because they are attracted by them but simply because they divide inside them. Smaller cycles inside the main cycles can also be seen. Those emerging global behaviours can also be observed in real predator-prey systems (see the Canadian lynx and snowshoe hare pelt-trading records of the Hudson Bay Company over almost a century (Odum, 1953)).

The simulation expresses classical cycles of predator-prey systems, indicating that the results are qualitatively correct.

Figure 8. Screen captures of the simulation at time steps 1490 and 12840. From left to right: the environment states, the population curves of entities Bdel (dark grey) and Leio (light grey) according to time, and the population curve of Leio according to Bdel (after change of the abscissa)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Water-Blood Dioxygen Exchanges in Zebra Fish

Real System

Zebra fishes have an efficient respiratory organ able to extract dioxygen from the water to their blood. Their gills use a counter-current system to improve the dioxygen exchange. Figure 9 represents how the system works. The water contains dioxygen molecules which circulate from the left to the right. A permeable membrane, that only gazes can cross (dotted line), separates the blood from the water. The dioxygen can cross the membrane from the water to the blood or from the blood to the water with exactly the same probability. In the real system, the blood circulates in the opposite direction. However, by using the simulation, the benefits of such mechanism will be studied as compared with the co-current one. The simulation will aim to evaluate the impact of the current direction (counter-current co-current) on the blood oxygenation.

Simplification

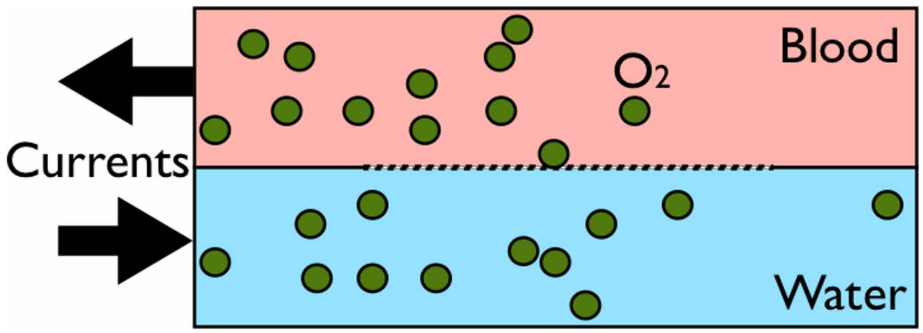
The simulation focuses on a small part of the system of the Zebra Fish, where the dioxygen exchanges occur. The red blood cells responsible for the transport of the dioxygen in the blood are not modelled, because this mechanism occurs essentially after the exchange. Other blood components, like immune cells for example, are not modelled because they do not interfere in the dioxygen exchange.

Model

The model is made to test the efficiency of the zebra fish counter-current respiratory system where the O_2 is central. The chosen model has three types of biological entities, the O_2 in various states, and two membranes:

- The O_2 is divided into three types of entity in order to give them different behaviours and count them according to their locations:
  - The O_2 in water (O2Water).
  - The O_2 in blood (O2Blood).

Figure 9. Mechanisms of dioxygen exchange inside the zebra fish gills



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

    - The O_2 in water which can no longer go to the blood (O2WaterLost) because it has been evacuated outside the respiratory system.
- The permeable membrane that can be crossed by the O_2 (Membrane).
- The impermeable membrane (Wall).

To complete the model, four dedicated types of entity have been used:

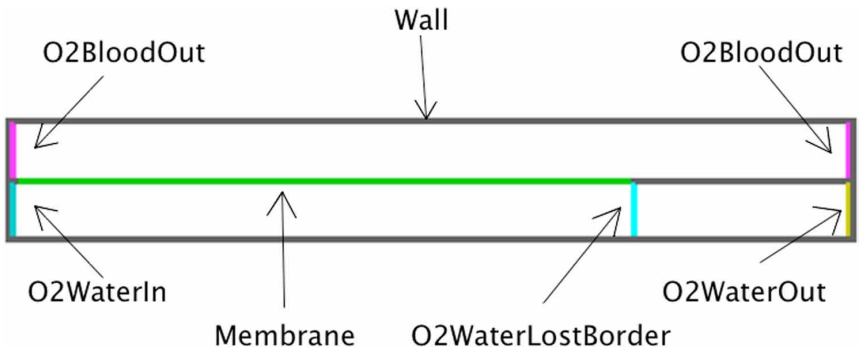
1. Entities O2WaterIn, which perform the addition of O2Water thanks to the behaviour called O2WaterGetIn.
2. Entities O2WaterOut, which remove O2Water thanks to the behaviour O2WaterGetOut.
3. Entities O2BloodOut, which remove O2Blood thanks to the behaviour O2BloodGetOut.
4. Entities O2WaterLostBorder, which detect O2Water entities that can no more go into the blood thanks to the behaviour O2WaterIsLost.

Figure 10 shows how entities are located in their grid environment at t=0.

Only two biological entities are placed: the impermeable membrane (Wall) and the permeable membrane (Membrane) which creates the upper row of blood circulation and the lower row of water circulation. There is also the four ad-hoc entities (non-biological entities). Note that the O_2 is not present at t=0 but will be added at the right to the ad-hoc entities O2WaterIn. When O_2 entities in the water diffuse to the right, they can cross the Membrane to go in the blood. The O_2 in the blood can also go back to the water by crossing again the Membrane. The O_2 entities in water reaching the O2WaterLostBorder can no longer cross the Membrane, and they become lost for the blood (O2WaterLost). The region between O2WaterLostBorder and O2WaterOut only contains O2WaterLost entities that will be counted to test the efficiency of the current and counter-current mechanisms. The O2BloodOut are located at both left and right sides of the upper row in order to evacuate the O_2 in the blood whatever the direction the current follows.

Nine behaviours have been made to add movement, membrane crossing, transformation, creation and removal of entities. O_2 entities move up, down or right at random in water at each simulation step (probability set to 1). The same mechanism is used to define the movements of O_2 in blood and O2WaterLost entities. Note that this behaviour can be modified by changing the left direction by the right, giving the

Figure 10. The gill model at its initial state (t=0)



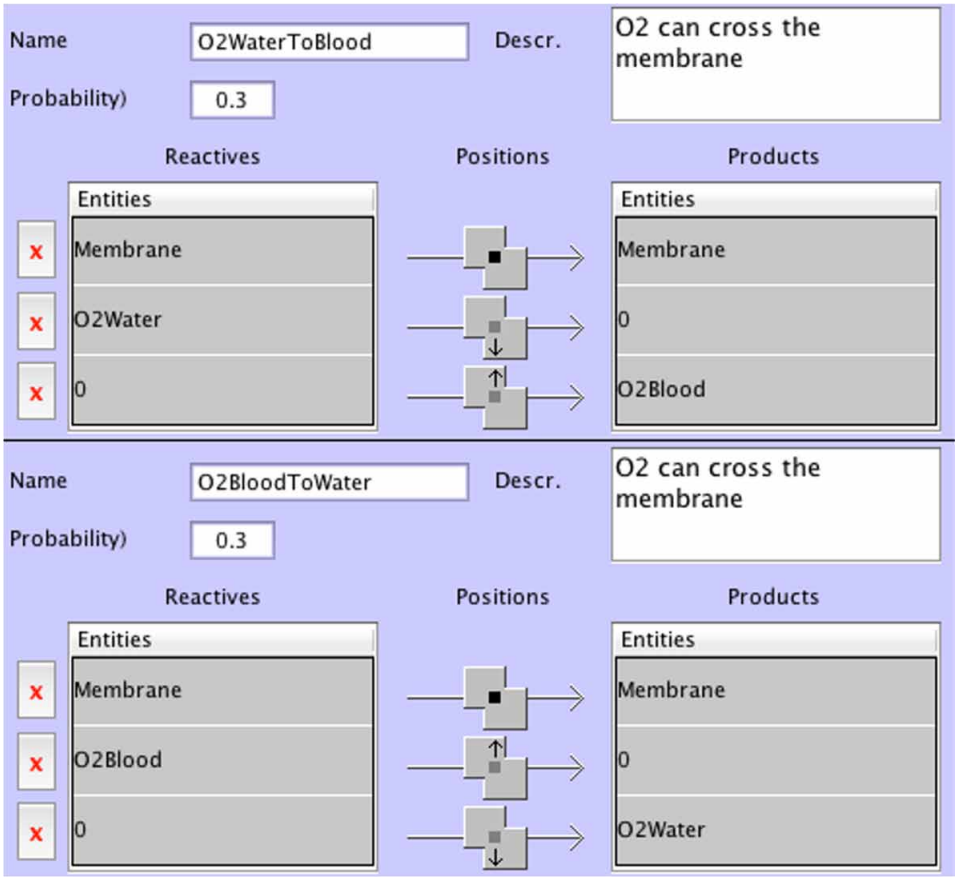
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

possibility to test the impact of the O_2 transfer efficiency from the water to the blood according to the relative direction of both flows.

Figure 11 shows how the O_2 crosses the membrane. The behaviour indicating how the O_2entities in water become lost for the blood is described in Figure 12.

Finally, Figure 13 shows the only one behaviour that adds entities in the environment. It creates O2Water entities just on the right of O2WaterIn entities with a probability of 0.1. It means that one time step over 10 (on average), one O2WaterIn adds one O2Water (if there is an empty space at its right). The same mechanisms are used to define two more behaviours: the behaviour removing the O_2 entities in the water that are too far from the membrane to be useful for the blood (O2WaterLost), and the behaviour removing the O_2 entities in blood (probabilities at 1). Defining these behaviours led to create dedicated entities O2WaterIn, O2WaterOut and O2BloodOut that do not exist in the real system. They only are here to simulate the continuous flow of O_2 in the water coming from the left and disappearing on the right, when they go out of the simulation area.

Figure 11. The two behaviours of this figure indicate how O_2 entities cross the membrane separating the blood and the water. The left one concerns the crossing of O_2 in water (O2Water) to the blood (O2Blood). The right one describes the crossing of O_2 located in blood (O2Blood) toward the water. Both of these behaviours have the same probability: 0.3



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. This behaviour changes O2Water entities touching the O2WaterLostBorder into O2Water-Lost entities. Thus, O_2 that cannot be used by the blood can be counted to evaluate the efficiency of the counter-current mechanism as compared with the current one. The less O2WaterLost entities are present, the more the gills are efficient

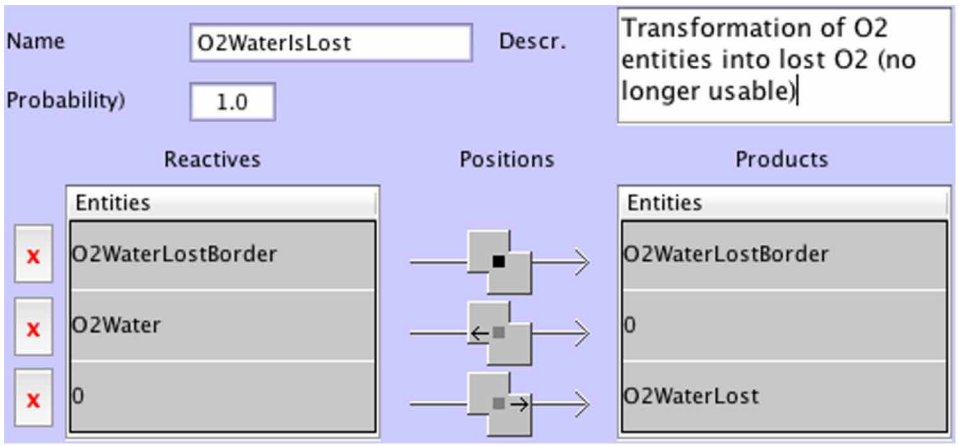
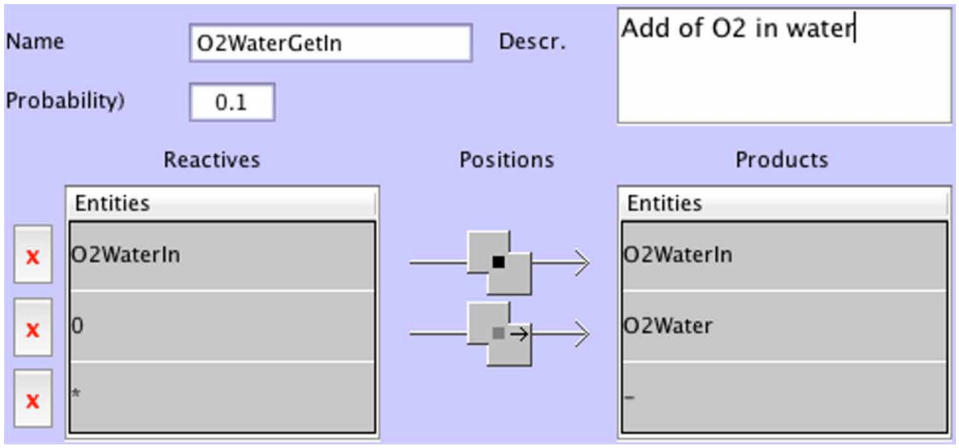


Figure 13. The behaviour adding O2Water entities in the environment



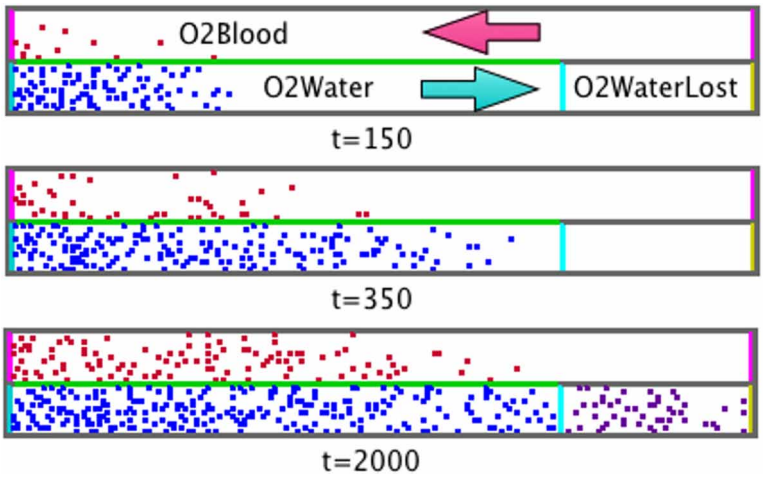
Simulation and Results

The simulation using the counter-current behaviour is shown as an example of the dynamics of the gills (see figure 14).

The counting of O2WaterLost gives the efficiency of the system and shows that, for the given initial state and the given behaviours' parameters, the co-current mechanism is about 42% worse in O2WaterLost as compared with the counter-current one. To evaluate the benefit of the counter-current direction as compared with the same current direction, the number of O2WaterLost has to be counted when the simulation is in a relative permanent state. To calculate when the permanent state arises, we need to know how long it takes for the O_2 newly created in water to cross the entire lower compartment. By looking at the behaviour of O2Water (normal or lost), it is clear that the right movement occurs, in

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Three screen captures show how the system evolves when the currents are in opposition (water comes from the left and circulates to the right; blood circulates from the right to the left). At t=150, the O2Water created by the O2WaterGetIn behaviour starts to diffuse to the right of the water compartment (the lower one). The O2Blood entities that reach the O2BloodOut are removed from the simulation thanks to the O2BloodGetOut behaviour. Few O2Blood, coming from the O2Water having crossed the membrane, move in the opposite direction. At t=350, O2Water entities reach the O2WaterLostBorder. At t=2000, the system is in a relatively stable state and O2WaterLost entities can be seen



average, one time over three (the two other possible movements are up and down). The length of the compartment is about 100. So, at least, 300 time steps are needed for the O_2 in the water to run through the lower compartment. Moreover, the O_2 can cross the membrane and go back to the left, following the blood flow. About 3/4 of the compartment have the permeable membrane, slowing down the movement of the O_2 in the water. Because this is a stochastic simulation, only probabilities can be calculated. The chosen criterion is arbitrary set by looking at the curves of O2LostWater concentration and by choosing a time from which the state looks permanent enough. The average number of O2LostWater is calculated between 1000 and 4000, which seem consistent (see Figure 15).

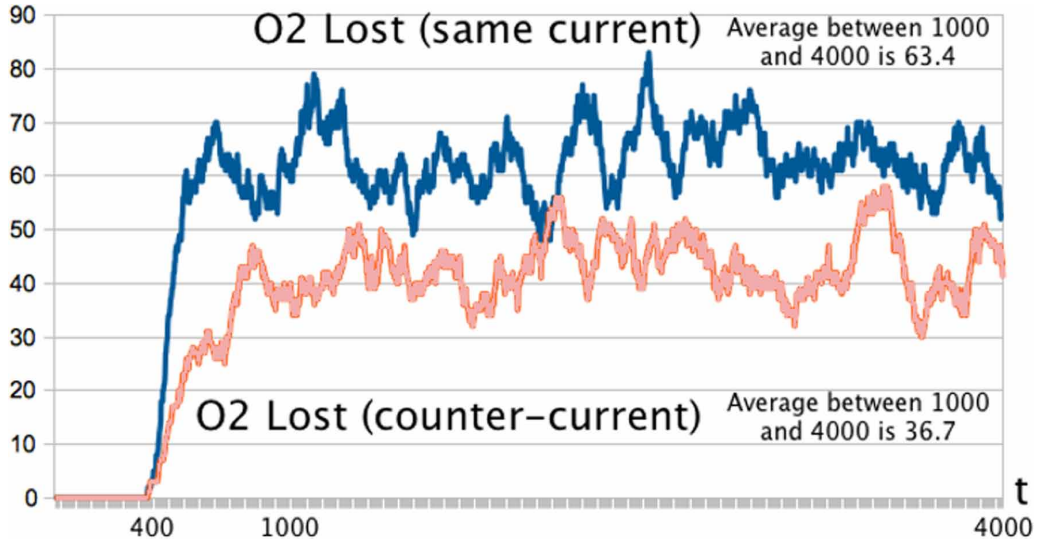
Blood Coagulation

Real System

The real system simulated in this example is a small section of a vein, in which a lesion is present. The aim of this model is to reproduce some of the molecular and cellular mechanisms during the blood coagulation leading to the formation of a blood clot. Three stages are usually described. The primary hemostasis starts with the vein lesion and finishes with the formation of a plug covering the lesion. The plug is made of platelets and soluble rod-like shaped proteins called fibrinogen. The secondary hemostasis is the transformation of the platelets plug into a blood clot thanks to the conversion of the fibrinogen into insoluble strands called fibrins. The third and final stage is the fibrinolysis, when the clot is broken down.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. The average number of O2WaterLost under the co-current condition is 63.4 and 36.7 under the counter-current condition. That makes a loss of O_2 which is 42% superior in the case of co-current. Note that this percentage depends on the size of the compartments and the size of the membrane: this can be easily seen on the screen captures where the density of O_2 in water decreases according to the distance from the O_2 creation, at the left



Model

The model presented here includes the primary hemostasis but not the secondary one nor the fibrinolysis. Moreover, not all the known mechanisms are simulated to improve the model clarity. The model is described using all the involved entities (cells, molecules and others), their initial state (stored into a 2D matrix) and their behaviours (interactions, transformations, creations and blood flow).

Two types of entities can be distinguished. The first ones are those which directly represent real biological entities like molecules or cells:

1. Fibro (for the fibroblast cells).
2. Endo (for the endothelial cells).
3. PlaC (for the circulating platelets).
4. PlaF (for the fixated platelets).
5. Fibri (for the fibrinogen molecule).

The second ones are those which do not exist as such, but are needed to simulate abstract biological mechanisms, like the blood flow in this case:

6. FlowS (for start of the blood flow).
7. FlowE (for end of the blood flow).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The initial state looks like a rectangle in which the upper edge consists of one layer of endothelial cells. The lower edge contains one layer of endothelial cells (with a discontinuity representing the lesion in the vein) plus one layer of fibroblast cells, below the endothelial layer. The left edge is made of entities regulating the blood flow entrance and the right side consists of entities regulating the blood flow exit (see Figure 16).

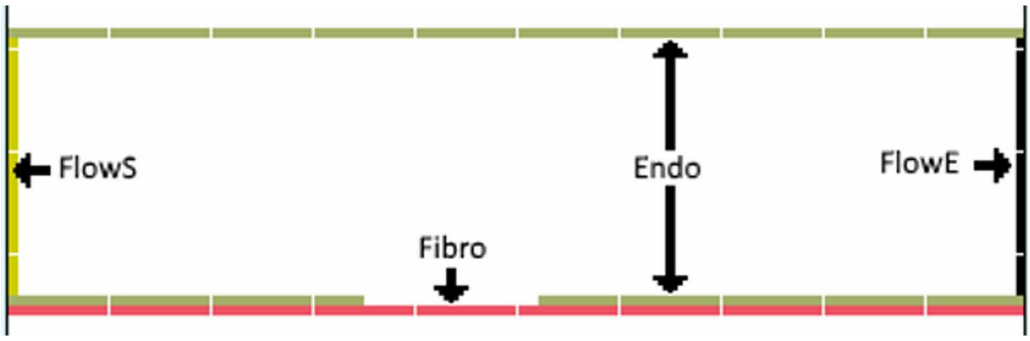
The behaviours developed in this simulation aim to model the blood flow and the core mechanisms of primary hemostasis. Firstly, five behaviours are needed to model the blood flow (see Figure 17):

- One behaviour to model the apparition of the platelets coming from the left of the vein section.
- Two behaviours to model the blood flow circulating from the left to the right which are applied to the platelets (PlaC entities) and the fibrinogens (Fibri entities).
- Two behaviours to model the disappearance of the circulating entities (PlaC and Fibri) when they reach the right of the vein section.

Secondly, three behaviours are needed to model the coagulation (see Figure 18):

- The first behaviour models the fixation of circulating platelets on the lesion (on the fibroblasts).
- The second one models the fixation of circulating platelets on the plug of platelets in presence of fibrinogens.
- The third behaviour models the indirect production of fibrinogens by fibroblasts. Note that this last behaviour is a simplification of a more complex molecular pathway.

Figure 16. The initial state, and more generally all the simulation, takes place into a 100x28 matrix where entities are distributed according to a simple rectangle corresponding to a vein section. The vein is made-up of three layers, two of endothelial cells (Endo entities) and one layer of fibroblast cells (Fibro entities). Note that the bottom layer of endothelial cells has a hole in order to simulate a lesion, resulting in a direct contact between the fibroblast cells and the circulating blood. The two vertical edges do not exist as real biological entities but are necessary to represent, at the left side of the vein, the incoming of circulating entities (FlowS entities) like the platelets and, at the right side of the vein, the disappearing entities (FlowE entities). The flow of circulating entities moves from the left (Start) to the right (End)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 17. The blood flow is simulated thanks to five behaviours. The two firsts (upper-left) are the movements of circulating platelets and fibrinogens entities from the left to the right. Note that they can also move vertically to reproduce the non-linearity of the blood flow. The third behaviour (upper-right) is the addition of platelets from FlowS (start of flow at the left of the vein section). The probability is 1% and can be changed to simulate different concentrations of platelets (according to the patient for example). The fourth and fifth ones (bottom of the figure) simply remove circulating entities that reach the end of the simulated vein section

Name	**PlatCMvt / FibriMvt**	
Description	Movement from the left to the right of the circulating platelets / fibrinogens	
Probability	1.0	
Reactives	Positions	Products
PlaC / Fibri	+	0
0	↑ PlaC / Fibri → ↓	PlaC / Fibri
*		-

Name	**AddPlaC**	
Description	"Creation"of circulating platelets from the start of the flow	
Probability	0.01	
Reactives	Positions	Products
FlowS	+	FlowS
0	FlowS →	PlaC
*		-

Name	**DelPlaC**	
Description	Removal of circulating platelets arriving at the right of the vein section	
Probability	1.0	
Reactives	Positions	Products
FlowE	+	FlowE
PlaC	← FlowE	0
*		-

Name	**DelFibri**	
Description	Removal of circulating fibrinogens arriving at the right of the vein section	
Probability	1.0	
Reactives	Positions	Products
FlowE	+	FlowE
Fibri	← FlowE	0
*		-

The end of the growth of the platelet plug is not modeled thanks to a behaviour but emerges from the simulation. More precisely, when all the fibroblast cells are covered with the anchored platelets, they simply cannot produce fibrinogen, stopping the growth of the plug.

Simulation and Results

In the simulation, 1 step corresponds to 1 second. The size of the vein is about 3 millimeter long and 1 millimeter large. As outputs, the shape of the plug is observed and the quantity of circulating fibrinogen is measured at each time step (see Figure 19).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. The core mechanism governing the apparition of a plug of platelets is only made with three behaviours. The first behaviour (upper-left) is the creation of fibrinogen entities from fibroblast ones. In the real system, it does not occur like this, but it is a convenient simplification for this simple example. The second one is the fixation of circulating platelets that enter in contact with the fibroblast cells of the lesion. The third one, involving three reactive entities, simulates the fixation of circulating platelets onto the plug of anchored platelets, allowing it to grow

Name	**AddFibri**	
Description	The fibroblasts *indirectly* create fibrinogens	
Probability	1.0	
Reactives	Positions	Products
Fibro	+	Fibro
0	↑ Fibro	Fibri
*		-

Name	**PlugBeginning**	
Description	Fixation of circulating platelets onto the fibroblasts in the lesion	
Probability	1.0	
Reactives	Positions	Products
PlaC	+	PlaF
Fibro	↑ ← PlaC → ↓	Fibro
*		-

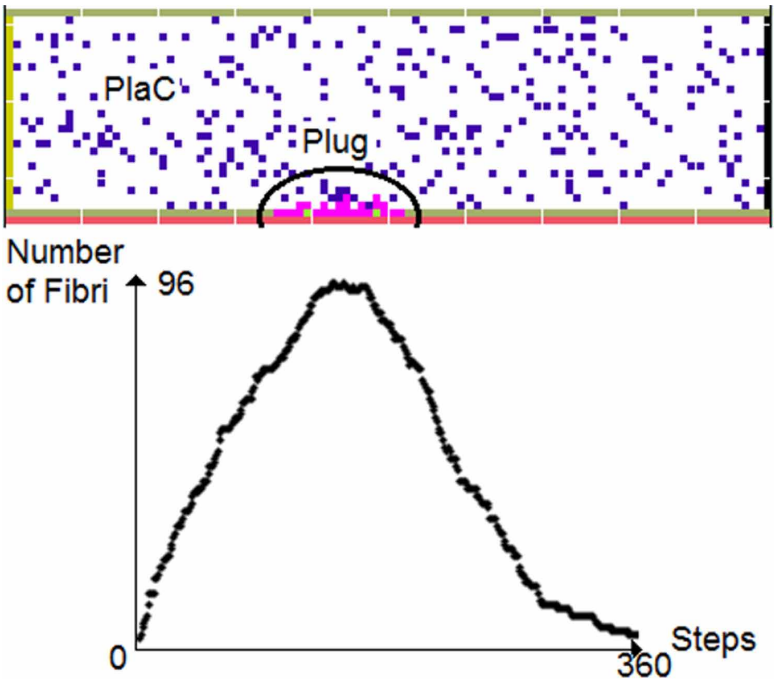
Name	**PlugGrowth**	
Description	Fixation of circulating platelets on the platelets plug in presence of fibrinogens	
Probability	1.0	
Reactives	Positions	Products
PlaC	+	PlaF
PlaF	↑ ← PlaC → ↓	PlaF
Fibri	↑ ← PlaC → ↓	0

The simulation shows that the plug of platelets starts after 139 steps (2 minutes and 19 seconds) and ends at step 254 (4 minutes and 14 seconds). After this step, no more fibrinogen entity is produced because the fibroblasts are no more in contact with the blood flow. The remaining fibrinogen disappears at the right of the vein section. The simulation is stopped at 360 steps, when all fibrinogen entities have disappeared.

With such a model, it could be possible to use the results of different simulations to build a synthetic and formalised mathematical model of the primary hemostasis. The virtual experiments can be performed quickly and at a very little price. Another interesting feature of such a model is that during a simulation it is possible to create new lesions, change the size or the shape of lesions, modify the platelet concentration, etc. The user can experiment many different scenario and hypothesis and see the validity but also the limits of the model.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 19. After 360 steps of simulation, corresponding to 6 minutes, the plug is formed and the circulating platelets continue to flow from the left to the right (upper diagram of the figure). The quantity of circulating fibrinogen is drawn on the lower part of the figure, showing its growth up to the value of 96 which could be multiplied by a given factor to match real experiments



DISCUSSION

Here, the authors discuss the interests and the drawbacks of NetBioDyn.

NetBioDyn Interests

The use of software in biology and education is interesting, both for teachers and students. It enables teachers to explain complex systems in an intuitive and pseudo-experimental way. It also helps students to create their own computational biological systems and see the important stages by creating a virtual system that rebuilds the real system. They can learn efficiently how the real system works.

From the teacher point of view, the modelling and simulation using such a software enable to:

- Create explicative and predictive simulations which can be easily handled by students
- Teach the students how to “program” complex systems with very little knowledge in computer sciences.
- Improve the learning curve of students.
- Stimulate students’ creativity.
- Quickly find out which students have assimilated the course and which have not.
- See where, in the system, a student is right or wrong.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

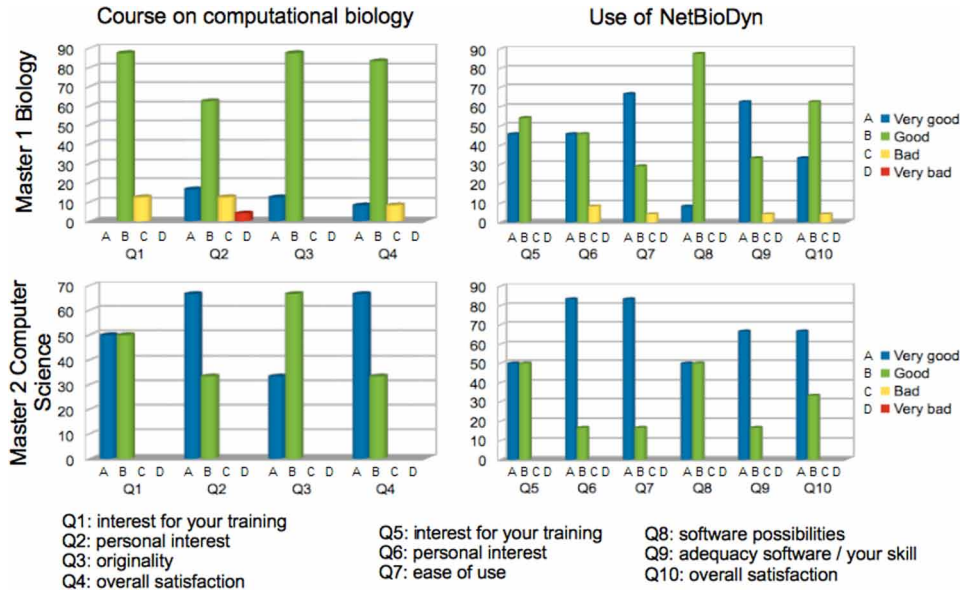
For the students, NetBioDyn gives the possibility to:

- Build or explore biological systems in an intuitive way with entities, behaviours and environments.
- Focus on the main mechanisms of a system and acquire a synthetic view.
- Pay attention to important mechanisms even if they look like details.
- Learn partially by themselves what is important and what is negligible.
- Organise their data and knowledge.
- Play at a motivating construction-like game.
- Explore the impact of parameters or behaviour variations in a simple way.
- Show the teacher what is their comprehension or misunderstanding of the real system.
- Share their knowledge with other students in a formal way.

A survey of 30 students (between 22 and 25 years old) has been made to assess the interest of the use of NetBioDyn in class. The results are presented in Figure 20. What is noteworthy is that biologist as well as computer scientist students are mostly interested by such an approach. The question of ease of use for NetBioDyn obtains the most "Very Good" answers for the two master students. Note that this survey is anonymous and that no teacher was present.

This software has also been successfully used by pupils (between 14 and 16 years old) of the middle school Max Jacob in Quimper (France) in May 2011.

Figure 20. Survey of 24 students in Master 1 Biology and 6 students in Master 2 Computer Science at the Université de Bretagne Occidentale. 10 questions have to be answered on the course they had on computational biology and on the use of the software NetBioDyn. Answers show that both the course and the software are considered important for their training. Moreover, the software is adapted to their skills (no code is required). Finally, the course on modelling and the software interest them and are useful for their studies



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

NetBioDyn Drawbacks

The limitations of NetBioDyn can be divided into two kinds. The first concerns the multi-agent approach itself and the second specifically concerns the software.

Multi-agent systems involve many interacting agents. This usually requires powerful computers, in term of computation and memory, to perform all the entities and their interactions. NetBioDyn reduces this problem by only implementing algorithms of complexity N+B (where N is the number of entities and B the number of behaviours). This is possible because the simulator only uses a small local neighbourhood on a grid. Another problem with multi-agent systems is the number of parameters that can be very important. The simulations made with NetBioDyn could become difficult to control and understand if too many types of entities and behaviours are added. Usually the authors recommend to their students not to add more than 20 types of entity or more than 20 different behaviours. Note that a type of entity is not an entity: a type of entity is a prototype that can be instantiated many times into entities without any problem of parameters.

Regarding the software limitations, the most important one is the entity shape that can only be a cube, implying that the simulations cannot be multi-scale. This last point can be partially circumvented as seen in all the presented examples where molecules, cells and tissues are mingled. For instance, an entity can represent far more than just one molecule, but a group of molecules. An entity can represent a single cell, but also a part of a tissue. What is important here is that an intrinsically multi-scale problem can generally be flattened.

CONCLUSION

The simulation of complex multi-cellular systems using computers represents a great challenge, both scientifically and technically. Scientifically, it serves the comprehension of fundamental mechanisms governing cells, from single-cell behaviours to whole systems involving numerous interacting cells.

Technically, the creation of computational models imitating multicellular systems requires advanced computational skills.

This is why the authors have developed and presented in this chapter a software making it accessible for non-computer specialists by simplifying the process of modelling and simulation.

This simulator is made to create and manipulate models whereas most of educational software are already-made models that researchers or students can test by only changing some parameters or initial state (see for example the simulator presented in (Meir, Perry, Stal, Maruca, & Klopfer, 2005)).

Three simulations have been presented to show the diversity of models that can be done and to explain how they can be made.

To support this statement, this software has been evaluated by pupils of middle school and higher education students, especially those in the field of biology.

The interest of this software has been discussed to bring out both positive contributions like focusing in what to model instead of how to model but also negative ones like the absence of multi-scale entities.

The perspectives are multiple and can be divided into two parts. Firstly the software itself can be improved. One of the most important features could be the possibility to factorise similar behaviours into one. This could be performed for the entities by adding generic behaviours like Random Walk. Another one is to increase the number of graphical tools to quickly draw shapes of entities like squares, circles

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

or filler. The authors are also working on a self-adjusting system to find the proper values for all the parameters involved in a simulation (behaviours' probabilities, entities half-lives etc.), in order to stay close to real results obtained for example *in vitro*.

Secondly, in terms of communication between NetBioDyn users, the main point to address will be the development of an internet social network where researchers, students and teachers will share their works, models and experiences (a starting professional network can be found at http://lefil.univ-brest.fr then NetBioDyn group).

REFERENCES

Aldrich, C. (2009). The Complete Guide to Simulations and Serious Games. John Wiley & Sons.

Ballet, P., Zemirline, A., & Marcé, L. (2004). The Biodyn language and simulators, application to an immune response and E. coli and phage interaction. In *Proceeding of the spring school on Modelling and simulation of biological processes in the context of genomics*.

Banitz, T., Gras, A., & Ginovart, M. (2015). Individual-based modeling of soil organic matter in NetLogo: Transparent, user-friendly, and open. *Environmental Modelling & Software*, *71*, 39–45. doi:10.1016/j.envsoft.2015.05.007

Bollen, L., & van Joolingen, W. R. (2013). Simsketch: Multiagent simulations based on learner-created sketches for early science education. *IEEE Transactions on Learning Technologies*, *6*(3), 208–216. doi:10.1109/TLT.2013.9

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., ... de Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics (Oxford, England)*, *25*(11), 1422–1423. doi:10.1093/bioinformatics/btp163 PMID:19304878

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, *4*(5), 41–60. doi:10.1002/(SICI)1099-0526(199905/06)4:5<41::AID-CPLX9>3.0.CO;2-F

Gammack, D. (2015). Using NetLogo as a tool to encourage scientific thinking across disciplines. *Journal of Teaching and Learning with Technology*, *4*(1), 22–39. doi:10.14434/jotlt.v4n1.12946

Ginovart, M. (2014). Discovering the power of individual-based modelling in teaching and learning: The study of a predator-prey system. *Journal of Science Education and Technology*, *23*(4), 496–513. doi:10.100710956-013-9480-6

Gkiolmas, A., Karamanos, K., Chalkidis, A., Skordoulis, C., & Papaconstantinou, M. D. S. (2013). Using simulation of Netlogo as a tool for introducing Greek high-school students to eco-systemic thinking. *Advances in Systems Science and Application*, *13*(3), 275–297.

Houser, N. (1992). *The Essential Peirce: Selected Philosophical Writings (1867-1893)* (Vol. 1). Indiana University Press.

Izaguirre, J. A., Chaturvedi, R., Huang, C., Cickovski, T., Coffland, J., Thomas, G., & Glazier, J. A. (2004). Compucell, a multi-model framework for simulation of morphogenesis. *Bioinformatics (Oxford, England)*, *20*(7), 1129–1137. doi:10.1093/bioinformatics/bth050 PMID:14764549

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning science. *Journal of the Learning Sciences*, *15*(1), 11–34. doi:10.120715327809jls1501_4

Jenkins, S. (2015). *Tools for critical thinking in biology*. New York, NY: Oxford University Press.

Kang, S., Kahan, S., McDermott, J., Flann, N., & Shmulevich, I. (2014). Biocellion: Accelerating computer simulation of multicellular biological system models. *Bioinformatics (Oxford, England)*, *30*(21), 3101–3108. doi:10.1093/bioinformatics/btu498 PMID:25064572

Kang, S., Kahan, S., & Momeni, B. (2014). Simulating microbial community patterning using biocellion. In Engineering and Analyzing Multicellular Systems (pp. 233-253). Springer. doi:10.1007/978-1-4939-0554-6_16

Klein, J. (2014). *Breve* [computer software]. Available from http://www.spiderland.org/breve/

Klein, J., & Spector, L. (2009). 3D Multi-Agent Simulations in the Breve Simulation Environment. In M. Komosinski & A. Adamatzky (Eds.), *Artificial Life Models in Software* (pp. 79–106). Springer. doi:10.1007/978-1-84882-285-6_4

Kubera, Y., Mathieu, P., & Picault, S. (2008). Interaction-oriented agent simulations: From theory to implementation. In *Proceedings of the 18th European Conference on Artificial Intelligence* (pp. 383-387). IOS Press.

Lynch, S. C., & Ferguson, J. (2014). Reasoning about complexity - software models as external representations. In *Proceedings of the 25th Workshop of The Psychology of Programming Interest Group*.

Meir, E., Perry, J., Stal, D., Maruca, S., & Klopfer, E. (2005). How effective are simulated molecular-level experiments for teaching diffusion and osmosis? *Cell Biology Education*, *4*(3), 235–248. doi:10.1187/cbe.04-09-0049 PMID:16220144

Merks, R. M. H., Guravage, M., Inzé, D., & Beemstaer, G. T. S. (2001). VirtualLeaf: An open-source framework for cell-based modelling of plant tissue growth and development. *Plant Physiology*, *155*(2), 656–666. doi:10.1104/pp.110.167619 PMID:21148415

Nikolai, C., & Madey, G. (2009). Tools of the trade: A survey of various agent based modelling platforms. *Journal of Artificial Societies and Social Simulation*, *12*(2), 2.

North, M. J., Collier, N. T., & Vos, J. R. (2006). Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit. *ACM Transactions on Modeling and Computer Simulation*, *16*(1), 1–25. doi:10.1145/1122012.1122013

Odum, E. P. (1953). Fundamentals of Ecology. Cengage Learning.

Prain, V., & Waldrip, B. (2006). An exploratory study of teachers and students use of multimodal representations of concepts in primary science. *International Journal of Science Education*, *28*(15), 1843–1866. doi:10.1080/09500690600718294

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Repenning, A., Smith, C., Owen, B., & Repenning, N. (2012). Agentcubes: Enabling 3d creativity by addressing cognitive and affective programming challenges. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 2762-2771). Chesapeake, VA: AACE.

Repenning, A., & Sumner, T. (1995). Agentsheets: A medium for creating domain-oriented visual languages. *IEEE Computer*, *28*(3), 17–25. doi:10.1109/2.366152

Schrödinger, L. L. C. (2015). *The PyMOL Molecular Graphics System, Version 1.8*. Available at http://pymol.org

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. doi:10.1101/gr.1239303 PMID:14597658

Shin, D. G., Liu, L., Loew, M., & Schaff, J. (1998). Virtual cell: A general framework for simulating and visualizing cellular physiology. *Visual Database Systems*, *4*, 214–220.

Stroup, W. M., & Wilensky, U. (2014). On the embedded complementary of agent-based and aggregate reasoning in students developing understanding of dynamic systems. *Technology. Knowledge and Learning*, *19*(1-2), 19–52. doi:10.100710758-014-9218-4

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, *31*, 455–461. PMID:19499576

Wallace, A. C., Laskowski, R. A., & Thornton, J. M. (1996). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering*, *8*(2), 127–134. doi:10.1093/protein/8.2.127 PMID:7630882

Weiss, G. (Ed.). (1999). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA: MIT Press.

Wilensky, U. (1999). *Netlogo*. Center for Connected Learning and Computer-Based Modeling. Available at http://ccl.northwestern.edu/netlogo/

XJ Technologies Company Ltd. (2012). *AnyLogic* [computer software]. Available from www.xjtek.com

Zhang, J., & Patel, V. L. (2006). Distributed cognition, representation, and affordance. *Pragmatics & Cognition*, *14*(2), 333–341. doi:10.1075/pc.14.2.12zha

KEY TERMS AND DEFINITIONS

Behaviour: A rule that transform or move entities in the simulated environment.

Bdellovibrio: A motile, vibrio shaped, monoflagellated, gram negative bacteria. It is a predator of Photobacterium leiognathi.

Entity: A virtual entity representing a cell or a molecule.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Integrated Development Environment (IDE): A software with an advanced graphical user interface (GUI) aiding the user to create other software applications.

Lysis: The death of an entity representing a cell.

Multi-Agent System: A set of interacting agents (also called entities). In our system, each agent has behaviours to perform simple reactive tasks.

Photobacterium Leiognathi: A bacteria which is bioluminescent and a prey for Bdellovibrio.

This research was previously published in Multi-Agent-Based Simulations Applied to Biological and Environmental Systems edited by Diana Francisca Adamatti; pages 128-158; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 49
Analysis of Electrocardiogram Data Compression Techniques: A MATLAB-Based Approach

Anukul Pandey
Dr. B. R. Ambedkar National Institute of Technology, India

Barjinder Singh Saini
Dr. B. R. Ambedkar National Institute of Technology, India

Butta Singh
Guru Nanak Dev University Regional Campus, India

Neetu Sood
Dr. B. R. Ambedkar National Institute of Technology, India

ABSTRACT

In this Chapter, a MATLAB-based approach is presented for compression of Electrocardiogram (ECG) data. The methodology employs in three different domains namely direct, transformed and parameter extraction methods. The selected techniques from direct ECG compression methods are TP, AZTEC, Fan, and Cortes. Moreover selected techniques from transformed ECG compression methods are Walsh Transform, DCT, and Wavelet transform. For each of the technique, the basic implementation of the algorithm was explored, and performance measures were calculated. All 48 records of MIT-BIH arrhythmia ECG database were employed for performance evaluation of various implemented techniques. Moreover, based on requirements, any basic techniques can be selected for further innovative processing that may include the lossless encoding.

DOI: 10.4018/978-1-7998-1204-3.ch049

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

In current era, patient monitoring systems are gaining importance which results in ever increasing data volume of various biomedical signals such as Electrocardiogram (ECG), Electroencephalogram (EEG) and Electromyogram (EMG). One of the open research problem in biomedical signal processing is the development of efficient data reduction technique (Sabarimalai Sur & Dandapat, 2014). Such techniques rely on maintaining a trade-off between minimization of reconstruction error and maximization of Compression Ratio (CR). ECG monitoring facilitates in providing faster health status of the concerned patient to the health care centre in case of the hostile cardiac behaviour. Cardiac monitoring using ECG signal is the best representative of heart electrical functionality and had proven to be effective in the diagnosis of most of the heart diseases. For continuous monitoring of patient's health status, the ECG signal is recorded for long time durations typically for several hours or few days. The storage of such huge volume of data requires large memory space, for example, a three-channel ECG signal sampled at a frequency of 1 kHz with 11 bits of resolution in three lead system of a particular patient, recorded for 24 h requires 928 MB of memory size per channel without any overhead. Such huge volume of data is also expected to have inter and intra beat correlation or inherent sparsity (Lee, Kim, & Lee, 2011; Ma, Zhang, & Dong, 2015).

So in the present chapter, a brief analysis of algorithms, mathematical background and its simulation implementation using MATLAB, for ECG data compression will be discussed. In order to make the ECG compression technique as a practical tool, this chapter is intended to refer to the MATLAB based routines. Matlab code (.m) of the most common ECG compression methods and algorithms are demonstrated in this chapter.

The performance evaluation of any ECG data compression techniques can be broadly assessed through following six parameters namely 1) Compression Ratio (CR), 2) Percent Root Mean Square Difference (PRD), 3) Percent Root Mean Square Difference Normalized (PRDN), 4) Root Mean Square (RMS) error, 5) Signal to Noise Ratio (SNR), 6) Quality Score (QS). These six qualitative performance evaluation parameters mentioned above are extensively used in literature for accessing the efficiency of any of the compression techniques.

The compression measure CR is equally applicable to both domains where there is difference or same sample resolution in compressed data. The degree of distortion measure in reconstructed ECG signal is primarily computed with respect to the original data, based on the visual as well the qualitative evaluation. The six qualitative performance evaluation parameters mentioned above are extensively used in literature for the comparison purpose (Batista, Melcher, & Carvalho, 2001; Blanco-Velasco et al., 2005; Lee et al., 2011).

The comparison of existing ECG compression methods using the above mentioned performance measures can be carried out, by performing experimentations over most cited Massachusetts Institute of Technology - Beth Israel Hospital (MIT-BIH) database.

Any digital data source has inherent redundancy, if the generated symbols are not equi-probable. The fundamental goal of any data compression approach is efficient data transmission or storage while preserving the significant diagnostic features. The process of data compression seeks to reduce the number of bits of information required to store or transmit digitized ECG signals without a significant reduction in the signal quality (ECG features). Data reduction of ECG signal is achieved by discarding digitized samples that are not vital for rhythm interpretation. For such data reduction, compression techniques are designed, which are classified as lossless or lossy techniques.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The lossless data compression guarantees the integrity of reconstructed data at a compromised CR, typically CR level of 2 to 1 with exact replication of original ECG while the lossy data compression achieves a high compression ratio with varying level of reconstruction error.

For ECG data compression, mostly the lossy type has been applied due to its capability of a high data compression ratio. The efficiency of any compression algorithm can only be evaluated after reconstructing the original signal through the process of decompression. The PRD is used as a similarity measure for the reconstructed and original ECG. The ECG data compression methods can also be broadly divided into three categories namely direct method, transformed method and parameter extraction method.

The direct data compression method analyzes and reduces data points in the time domain. Various direct data ECG compression schemes which includes Turning Point (TP), Amplitude Zone Time Epoch Coding (AZTEC), improved modified AZTEC technique, Co-ordinate Reduction Time Encoding System (CORTES), the delta algorithm and the Fan algorithm. They all have shown their ability for the efficient ECG data reduction.

Typically, AZTEC algorithm has the advantage of high data reduction ratio but suffers from loss of data fidelity due to the presence of discontinuity that occurs in the reconstructed ECG waveform. ASCII based encoding methods are also employed for compression due to its simplicity, which enables them to be used in portable and mobile data ECG data monitoring systems.

The transformed method analyzes energy distribution of the ECG samples by converting the time domain to some other domain, like Fourier transform, Fourier descriptor, Karhunen–Loeve Transform (KLT), the Walsh transform, the Discrete Cosine Transform (DCT) and the wavelet transform etc. This methods falls in category of lossy ECG compression methods as it mostly employs thresholding at some level. The main parameters which needs to be optimized are namely, CR, PRD and threshold value. The transformed based ECG compression method generally consist of following stages, i) pre-processing; for noise removal, ii) transformation; for visualization in other domain, iii) retention of significant coefficients, iv) application of quantization and v) lossless encoding. In the proposed chapter each and every stage was mathematically analyzed and the simulation results has been produced with MATLAB.

The parameter extraction method is based upon dominant feature extraction from the raw signal. The extracted parameter are quantized and then entropy coder is applied for enhancing the compression efficiency. This method includes the processes like vector quantization, peak picking and linear prediction method. There is scope of combining two or more ECG data compression techniques to get better compression efficiency in one or other domain which is an open research problem.

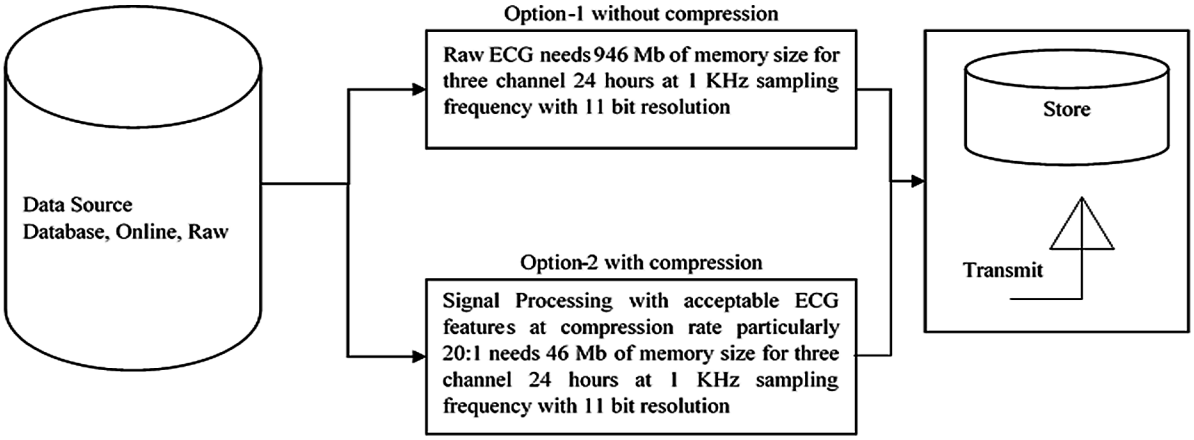
An extensive survey of the various compression methods highlighting their advantages, disadvantages, mathematical analysis and simulation results using MATLAB platform is presented in this chapter proposal thereby facilitating the future research in this direction.

Need of ECG Data Compression

In brief, ECG data compression techniques is the art of signal processing preserving the diagnostic information in a compact form. For any individual under cardiac observation, the ECG data is retrieved continuously online or offline as per prescribed by the cardiac expert. The digitation of ECG data is carried through at the acquisition stage by selecting appropriate sampling frequency and quantization levels. For example as shown in Figure 1, a three-channel ECG signal sampled at a frequency of 1 kHz with 11 bits of resolution in three lead system of a particular patient, recorded for 24 h requires 928 MB of memory size per channel without any overhead. Such huge volume of data is expected to have inter

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. ECG data memory size requirement with and without compression technique



and intra beat correlation or inherent sparsity. For monitoring purpose three lead configuration for ECG acquisition is appropriate. In general for clinical acceptability 12 lead configuration is recommended. This voluminous data can only be handled via appropriate compression technique with varying level of compression ratio and degree of distortion.

Redundancy in the ECG Data

ECG data compression is possible, because ECG, in general, are highly correlated in inter and intra beat which leads to redundancy. High rate of compression can be achieved through reduction in redundancy and correlation present in ECG data. Redundancy can be inferred as repeated intra beat data in ECG as quasi-stationary sense and higher statistical correlation in inter beat ECG represent plateaus.

DATABASE AND PERFORMANCE MEASURE

In the emerging field of medical engineering, research subjects like cardiac arrhythmia detection (Baibars, Kanjwal, & Marine, 2014), heart rate variability (Singh & Singh, 2012), cardiovascular and pulmonary dynamics (Cheng, 2015), artificial intelligence based medical decision support (Hebert, Korabek, & Scott, 2006)and ECG data compression(Jalaleddine, Hutchens, Strattan, & Coberly, 1990; Ma et al., 2015; Sabarimalai Sur & Dandapat, 2014), etc. are of major interest. The Massachusetts Institute of Technology (MIT) supplies some invaluable resources for such research projects. These resources include databases containing recorded physiological signals and software for analyzing, viewing and creating such recordings. In the present study, ECG samples from Arrhythmia Database (Goldberger et al., 2000) (http://ecg.mit.edu) have been taken for investigation of the efficiency of the various methods. The Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory. The 11-bit resolutions over 10 mV range with a sampling frequency of 360 Hz per sample channel were recorded. To access the compressor efficiency, first 21,600 samples (10 min duration) were used and evaluated with following six data objective compression performance measure metric (Batista et al., 2001; Blanco-Velasco et al., 2005; Lee et al., 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The .mat file can be extracted from (http://www.physionet.org/cgi-bin/atm/ATM) by selecting appropriate database which is MIT-BIH Arrhythmia Database (mitdb). The record number and appropriate time duration can be selected for export signal as .mat file from toolbox provided by website. For typical record 100, 117 and 119 the extracted transposed sample values for different leads are given in Table 1. The total 6, 50,000 ECG samples per lead is presented per record for 30 minutes of duration at 360 Hz sampling frequency. This total number of samples can be calculated approximately as 360×60×30.09 (sampling frequency× second to minute unit conversion factor × time in minutes).

The sample values stated in Table 1 are the benchmark for illustrating different preceding examples in this chapter. The performance measures are

Compression Ratio (CR)

The CR provides information about the degree by which the compression algorithm removes the redundant data. Higher the *CR* less number of bits required to store or transmits the data that can be defined as

$$CR = \frac{B_0}{B_c} \tag{1}$$

where, B_0 is the total number of bits required to represent original data and B_c total number of bits required to represent compressed data along with the side information needed for retrieval of original data.

Percentage Root Mean Square Difference (PRD)

It is a measure of acceptable fidelity and degree of distortion introduced during compression and decompression algorithm. Usually, the clinical acceptability of reconstructed signal after decompression is determined through visual inspection. The numerical evaluation of the level of distortion in reconstructed signal in comparison to original signal can be carried out via PRD measure, defined as

$$PRD(\%) = 100 \times \sqrt{\frac{\sum_{n=1}^{N} (X_s(n) - X_r(n))^2}{\sum_{n=1}^{N} (X_s(n))^2}} \tag{2}$$

where X_r is reconstructed signal with respect to original signal X_s.

Percentage Root Mean Square Difference, Normalized (PRDN)

It is a normalized version of PRD, is PRDN, which does not depend on the mean signal value $\overline{X}$. For any signal, the information lies in its variance. The mean value of the data does not actively contribute to the reconstructed signal fidelity issue. PRDN provides better error estimation in reconstructed signal than PRD.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Extracted discrete sample values for selected records from MIT-BIH Arrhythmia Database

	Record					
	100		117		119	
Sample Number	MLII	V5	MLII	V2	MLII	V1
1	995	1011	839	930	825	930
2	995	1011	839	930	825	930
3	995	1011	839	930	825	930
4	995	1011	839	930	825	930
5	995	1011	839	930	825	930
6	995	1011	839	930	825	930
7	995	1011	839	930	825	930
8	995	1011	840	930	821	930
9	1000	1008	840	930	827	933
10	997	1008	836	927	827	934
11	995	1007	836	926	830	935
12	994	1007	833	922	829	942
13	992	1009	830	921	830	939
14	993	1010	830	921	826	941
15	992	1010	828	921	824	940
16	989	1011	828	923	820	939
17	988	1013	831	924	819	935
18	987	1014	834	923	818	934
19	990	1014	836	928	815	933
20	993	1016	842	929	819	935
-	-	-	-	-	-	-
100	957	989	833	910	945	943
101	958	985	835	917	940	941
102	957	987	836	913	935	942
-	-	-	-	-	-	-
498	960	985	839	903	1236	735
499	961	987	840	906	1276	690
500	963	990	840	903	1312	650
-	-	-	-	-	-	-
650000	768	1024	1024	1024	1280	768

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$PRDN(\%) = 100 \times \sqrt{\frac{\sum_{n=1}^{N}(X_s(n) - X_r(n))^2}{\sum_{n=1}^{N}(X_s(n) - \overline{X})^2}} \tag{3}$$

Percentage Root Mean Square Difference, Normalized with Respect to Database (PRD_1024)

As for storage purpose, MIT-BIH database adds a 1024 in the original signal. So in order to evaluate the fidelity of reconstructed signal 1024 must be subtracted from the original signal.

$$PRD_1024 = 100 \times \sqrt{\frac{\sum_{n=1}^{N}(X_s(n) - X_r(n))^2}{\sum_{n=1}^{N}(X_s(n) - 1024)^2}} \tag{4}$$

Root Mean Square Error (RMS)

It provides a measure of error in a reconstructed signal with respect to original signal having signal length *N*.

$$RMS = \sqrt{\frac{\sum_{n=1}^{N}(X_s(n) - X_r(n))^2}{(N-1)}} \tag{5}$$

Signal to Noise Ratio (SNR)

SNR, measure of degree of noise energy introduced by compression algorithm in decibel (dB) scale.

$$SNR = 10 \times \log\left(\frac{\sum_{n=1}^{N}(X_s(n) - \overline{X})^2}{\sum_{n=1}^{N}(X_s(n) - X_r(n))^2}\right) \tag{6}$$

Quality Score (QS)

QS is defined as the ratio between the CR and PRD. The QS is a very rational performance indicator when it is difficult to estimate the best compression method while taking into account the reconstruction errors as well. A high score represents a superior compression method.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$QS = \frac{CR}{PRD} \tag{7}$$

All the seven measures are simultaneously used in literature for accessing the efficacy of any ECG compression algorithm. The MATLAB script for calculation for above mentioned performance metric are as follows, which will be used in this chapter for all compression techniques.

If gain (200) and base (1024) are removed before ECG signal processing, then before metric computation both gain and base should be incorporated in reconstructed ECG. Thereby making a fair comparison in the same condition with the original ECG.

Performance Metric Computation

```
ecg=ecg*200+1024;
ecg1=ecg1*200+1024;      % reconstructed ECG signal

l=length(ecg1);           % Length of reconstructed ECG signal
err=(ecg1-ecg(1:l)).^2;   % Square error computation between reconstructed ECG
and original ECG
nu=sum(err);              % Summation of square error

de_1024=(1024-ecg(1:l)).^2;   % Denominator for PRD_1024 computation
du=sum(de_1024);              % Summation of Denominator for PRD_1024
PRD_1024=100*sqrt(nu/du);

de=(ecg(1:l)).^2;         % Denominator for PRD computation
du=sum(de);               % Summation of Denominator for PRD

PRD=100*sqrt(nu/du);

de_mean=(ecg(1:l)-mean(ecg)).^2;  % Denominator for PRDN computation
du=sum(de_mean);                  % Summation of Denominator for PRDN
PRDN=100*sqrt(nu/du);

RMS=sqrt(nu/(length(err)-1));    % RMS computation

SNR=10*log10(du/nu);               % SNR computation
bit_resolution_original=11;     % Original sample resolution
bit_resolution_compressed=11;   % Compressed sample resolution
cr=(bit_resolution_original*length(ecg1))/(bit_resolution_
compressed*length(ecg_tp));

QS=cr/PRD;                             % QS computation
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Published with MATLAB® R2015a

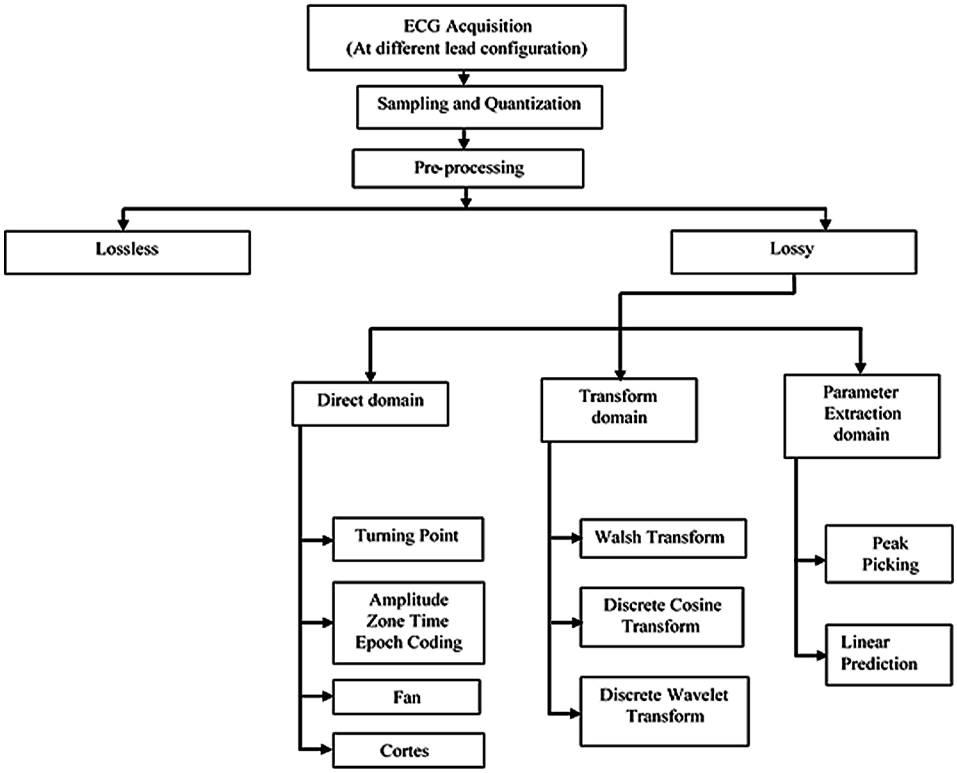
ECG Data Compression Methods

ECG data compression is broadly classified into two types: (i) lossless ECG data compression and (ii) lossy ECG data compression scheme. Lossless ECG data compression is preferred in case of critically ill patient monitoring where each subtle features are needed to be reproduced. Lossy ECG data compression is preferred in case of non-critical patient monitoring conditions and the elderly or regular monitoring purposes. In Lossless ECG compression exact replica of the original signal can be reproduced. But the achieved compression ratio is less than its counterpart i.e. lossy compression scheme. The achieved higher compression ratio in case of lossy ECG compression is at the cost of quality of reconstructed ECG signal. Figure 2, represents the flow chart giving a diversified classification of ECG data compression methods.

Lossless ECG Data Compression Methods

In lossless ECG data compression, the ECG after decompression and compression is identical to original ECG. Every bit of features and other information in reconstructed ECG is preserved. In other words we can state that the reconstructed ECG is exact replica of the original signal. Although this method provides the one of the conflicting requirements of compression, but fails to achieve the higher compression ratio.

Figure 2. Broad classification scheme for ECG data compression methods



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Advantage

- Higher reconstruction signal fidelity.
- Exact replication of original signal is possible.
- Reversible compression.

Disadvantage

- Lower compression ratio.
- Not applicable.

Application

- Critical patient monitoring.

Lossless ECG data compression technique (Horspool & Windels, 1994; Koski, 1997; Sayood, 2010) includes

1. Huffman Encoding
2. LZ Encoding
3. DPCM Encoding

Lossy ECG Data Compression Methods

In lossy compression method reconstructed ECG signal contains degradation with respect to original ECG signal. In this methodology of signal processing perfect reconstruction is sacrificed by the elimination of some amount of redundancies in the original ECG data to in order to achieve a higher compression ratio. Compared to lossless ECG compression method a higher compression ratio can be achieved with varying reconstruction error. For subjective evaluation of the compression scheme these reconstruction errors are not significantly visible. So, lossy ECG data compression methods under certain conditions are regarded as '*visually lossless'*. These certain conditions are preserving the diagnostic features, maintaining the base line. Depending upon the domain of operation on signal the lossy ECG data compression is further subdivided into three groups.

THREE METHODS FOR ECG DATA COMPRESSION

ECG data compression methods can be divided into three different groups namely i) Direct method ii) Transformed domain and iii) Parameter extraction method.

Direct Method

In direct ECG data compression method the digitized ECG samples directly analyzes and compresses the ECG data in the time domain. This section presents a brief overview of the each of basic algorithm

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

along with simulation mechanism especially for ECG data compression. These includes TP, AZTEC, CORTES and Fan algorithm.

Turning Point Method

The basic idea in TP method for ECG data reduction is *to reduce the sampling frequency* of original ECG signal by a factor of two. The maximum possible CR that can be obtained using the TP method is '2' because it preserves one of two samples retaining the slope/turning point information.

The algorithm process three ECG samples at a time. Considering the first sample as reference and then retain one of two remaining samples. The retained sample will become the reference for next two samples. The retention criteria of sample depends upon the magnitude of slope of two samples with respect to reference sample. Whichever sample out of two is having higher magnitude of slope that particular sample is retained other will be dropped for further processing. This process iterates till last sample of the ECG signal is encountered.

In past research (Mueller, 1977) author have concluded that at sampling frequency 200 Hz for ECG data, with 12 bit /sample resolution CR and PRD equal to 2 and 5.3 respectively were achieved. The basic TP procedure for compression is described in Algorithm 1, and reconstruction of ECG data is presented in Algorithm 2.

Algorithm 1. TP method for ECG data compression

```
1:        INPUT:
2:        ecg is the original ECG signal
3:        OUTPUT:
4:        ecg_tp is the compressed ECG signal
5:        Begin
6:        X_ref=ecg(1);
7:        ecg_tp(1)=x_ref;                          // Taking the first as
reference point
8:        For  i=2:2:length(ecg)-1
// Increment the loop index by 2 till the length of
the ecg signal
9:        Compute the slope of two point with respect to reference sample
10:        Retain the sample whose slope magnitude is higher with respect to
reference sample
11:        Retained sample become the reference sample for next iteration
12:        Concatenate the retained sample in ecg_tp
13:        End
```

Algorithm 2. TP method for reconstruction of ECG data

```
1:        INPUT:
2:        ecg_tp is the compressed ECG signal
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
3:         OUTPUTS:
4:         ecg1 estimate of original signal
5:         Spline interpolation (which inserts one point in between two consec-
utive compressed ECG samples)
6:         End
```

Matlab Simulation Program

Contents

- ECG Data Acquisition
- Initialization
- TP Method for ECG Data Compression
- Reconstruction
- End of TP Method for ECG Direct Data Compression Method

```
clc
clear
close all
```

ECG Data Acquisition

```
load('119m.mat');
fs = 360;
time_sec=30*60;                          % Time in number of seconds
ecg=(val(1,1:fs*time_sec)-1024)/200;  % Removing the base(1024) and gain(200)
```

Initialization

Initialization of variable ecg_tp, x_ref and ecg_tp(1)

```
ecg_tp=zeros(1,floor(length(ecg)/2));
x_ref=ecg(1);
ecg_tp(1)=x_ref;
```

TP Method for ECG Data Compression

```
j=2;
for i=2:2:length(ecg)-1
    % Selecting the pair of samples
    x1=ecg(i);
    x2=ecg(i+1);

    s1=x1-x_ref;
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
    s2=x2-x_ref;
    if abs(s1)>abs(s2) % Comparing the slope
        x_ref=x1;
    else
        x_ref=x2;
    end
    ecg_tp(j)=x_ref; % Updating the compressed ECG by selecting appropriate
data point
    j=j+1;
end
```

Reconstruction

```
ecg1=spline(1:length(ecg_tp),ecg_tp,1:.5:length(ecg_tp));
```

End of .m Script

An illustration for typical record 119 is given in Figure 3 which shows the original, reconstructed and error signal for 2000 samples. The above mentioned matlab script (.m) files is executed for all 48 records from MIT-BIH database. Table 2 shows the performance measure for typical record 100, 117, and 119 along with the average value for all 48 records. It can be noted that a constant CR equal to 2 and varying values of other performance metrics were obtained using this algorithm. These varying parameter results from the changing complexity in the signal and may violate the nyquist theorem in piece-wise sense.

Advantage

- It retains maximum and minimum points at which the sign of the slope of the signal changes.
- Simple and fast ECG compression algorithm to produce fixed CR of 2:1.
- Iterative application of TP method produces dyadic multiple CR, if there is much higher sampling rate the distortion error may be acceptable.

Limitations

- May violate sampling theorem criteria which results higher degree of distortion in reconstructed signal after repeated application of TP method.
- The retained samples are not equally spaced which may result in localized distortion.

Summary

TP method retains one of two samples through the maximum slope criteria with respect to reference sample point. The algorithm can provide fixed dyadic CR with varying values of reconstruction error. The reconstruction can be accomplished via interpolation between two consecutive samples.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Original, compressed, reconstructed and error ECG signal for TP algorithm for typical record 119

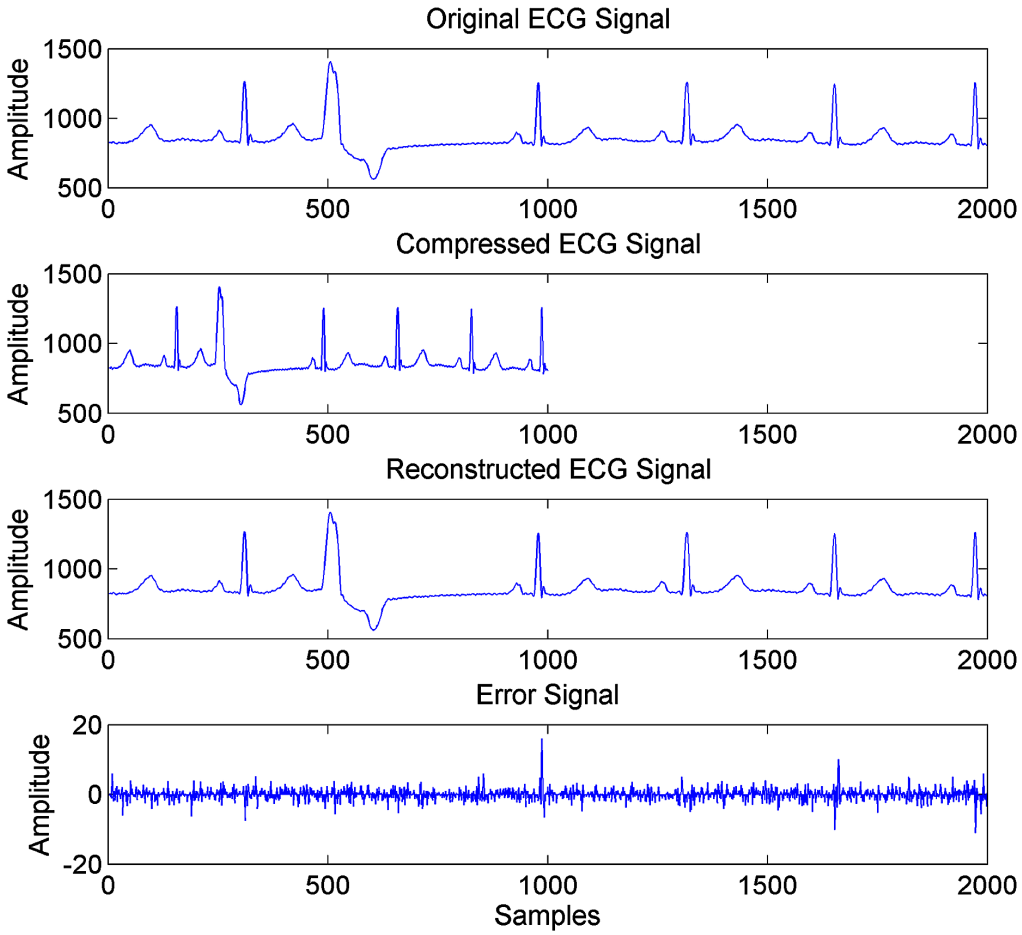


Table 2. Performance measure for selected and averaged value for all 48 records from MIT-BIH database

Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	2	1.79	0.13	3.36	1.3	29.47	14.83
117	2	0.77	0.16	2.85	1.36	30.91	12.57
119	2	0.72	0.17	1.37	1.46	37.3	11.74
A*	2±0	1.66±0.75	0.15±0.04	2.28±0.94	1.46±0.35	33.53±3.5	13.81±2.84

A*- for all 48 records average± standard deviation

AZTEC

The AZTEC is original developed for rhythmic analysis in real time monitoring of ECG. The basic AZTEC algorithm was first proposed by authors of (Cox, Nolle, Fozzard, & Oliver, 1968) which further improved modified in (V. Kumar, Saxena, & Giri, 2006). In present chapter we provide step by step il-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

lustration of basic AZTEC algorithm with its matlab implementation. The AZTEC algorithm decomposes the signal in terms of plateaus, extremum and slopes. The plateaus and slope detection thresholds are determined empirically. The Algorithm 3 shows the AZTEC algorithm for direct ECG data compression. The combination of plateaus and extremum is called bound.

Algorithm 3. AZTEC algorithm for direct ECG data compression

INPUT:
ecg is the original ECG signal
OUTPUT:
AZTEC is the compressed ECG data
Begin
Plateau detection
Compare the ECG signal sample with previous extrema point and experimentally determined threshold amplitude.
Count the detected consecutive sample which satisfies the above condition
Replace the samples with average value of extrema and name it as Line (L)
Slope detection
Application of experimentally determined threshold for estimating length of plateau
Check for the criteria of the extrema point using the following condition

$$\left\{\left(V(L_{i+1})-V(L_i)\right)\left(V(L_i)-V(L_{i-1})\right)\right\}<0$$

The combined extrema with plateau is called as bound. Except bound left over sample points will fall into the category of slope.
The concept of the time list of slope was devised by (Cox et al., 1968) and is given as

$$\tau(S)=\sum_{i=1}^{m-1}\tau(L_i)$$

The amplitude of slope is determined by

$V(S) = V(L_m) - V(L_0)$

where L_m and L_0 are previous bounds prior to slope.
In order to distinguish the slope and bound, we assign –ve sign to slope time list.
The complete AZTEC can be determined as an ordered pair of time lists and amplitude list.
End

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Matlab Simulation Program

Contents

- Plateau Detection
- Slope Detection
- Reconstruction of ECG signal
- plot
- Performance measure

```
%AZTEC COMPRESSION ON ECG
close all;
clear all;
clc;
load('119m.mat');
ecg=(val(1,1:2077)-1024)/200';
fs = 360;
a=length(ecg);
time = [1: a]/fs;
ln=0;i=1;j=1;tau=0;
vmin=ecg(1);vmax=ecg(1);
vmin1=vmin; vmax1=vmax;
tau_count=1;
```

Plateau Detection

```
for i=2:a

    if ecg(i)>vmax              % maximum and minimum value updatation
        vmax1=ecg(i);
    elseif ecg(i)<vmin
        vmin1=ecg(i);
    end
    %     avg_L=(vmax+vmin)/2;
    if vmax1-vmin1<=.1           % Check for experimentally determined thresh-
old
        tau_count=tau_count+1;
        avg=(vmax+vmin)/2;
        tau(j)=tau_count;               % duration of line information
        ecg_avg(j)=(vmax+vmin)/2;       % value of the line information
    else
        j=j+1;
        tau_count=1;
        tau(j)=1;
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
        ecg_avg(j)=ecg(i);

        vmin=ecg(i);vmax=ecg(i);
        vmin1=vmin; vmax1=vmax;
    end
end
loc=find(tau==0);
tau(loc)=1;
```

Slope Detection

A set of AZTEC lines is said to be a “slope” if no line Li is a bound and Lo and Lm are bounds. The slope S is completely specified by its value, v(S)=v(Lm)-v(Lo)

```
j=1;
ecg_all=[];
tau_all=[];
tau_all(j)=tau(j);
ecg_all(j)=ecg_avg(j);
oho=0;
for i=1:length(tau)-1
    if tau(i)<6        % confirmation of not being plaetu

        if  (ecg_avg(i+1)-ecg_avg(i))*(ecg_avg(i)-ecg_avg(i-1))<0% condition
for extrema detection

            if sum(oho)~=0
                tau_all(j)=sum(oho);
                j=j+1;
                oho=0;
            end
            tau_all(j)=tau(i);
            ecg_all(j)=ecg_avg(i);
            j=j+1;

        else
            % now slope operation
            oho=-(abs(oho)+ tau(i)); % keep eye on adding
        end
    else
        % plaetu
        if sum(oho)~=0
            tau_all(j)=sum(oho);
            j=j+1;
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
            oho=0;
        end
        tau_all(j)=tau(i);
        ecg_all(j)=ecg_avg(i);
        j=j+1;
    end
end

tau_all=tau_all'; % time list
tau=tau';
ecg_all=ecg_all';
% search for position of slopes
ind=find(tau_all<0);
ecg_all(ind)=ecg_all(ind+1)-ecg_all(ind-1);   % amplitude list

AZTEC(1:2:2*length(tau_all))=tau_all;
AZTEC(2:2:2*length(ecg_all))=ecg_all;
```

Reconstruction of ECG Signal

```
% reverse of slope
AZTEC;
tau_all1=AZTEC(1:2:length(AZTEC)); % should be same as tau_all
ecg_all1=AZTEC(2:2:length(AZTEC));
loc=find(tau_all1<0);
x=ecg_all1;
% inverse of slope
out=x(1:loc(1)-1);
tau1=tau_all1(1:loc(1)-1);
for i=1:length(loc)  % where to insert the slope information
    avg=x(loc(i))/(abs(tau_all1(loc(i)))+1);
    b=x(loc(i)-1)+avg:avg:x(loc(i)+1)-avg;
    % numbers to be inserted

    if loc(i)==loc(end)
        out=[out b x(loc(i)+1:loc(end)) x(end)];
        tau1=[tau1 ones(1,abs(abs(tau_all1(loc(i))))) tau_all1(end)];
    else
        out=[out b x(loc(i)+1:loc(i+1)-1)];
      tau1=[tau1 ones(1,abs(abs(tau_all1(loc(i))))) tau_
all1(loc(i)+1:loc(i+1)-1)];
    end
end
ecg1=[];
```

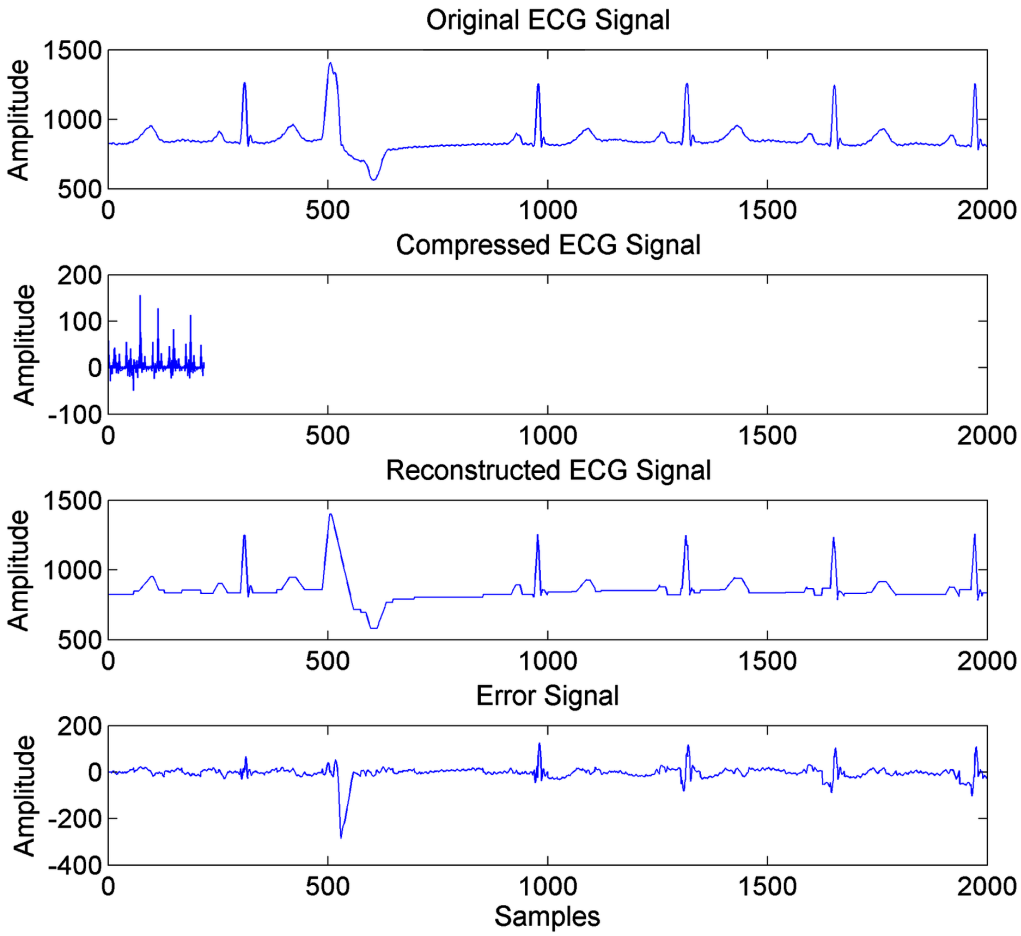

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
for i=1:length(out)
    ecg1=[ecg1 rectpulse(out(i),tau1(i))']; % few ending samples might be
missing
end
```

End of .m Script

Figure 4, the original, reconstructed and error signal is depicted for 2000 samples. The above mentioned matlab script (.m) file is executed for all 48 records from MIT-BIH database and results are presented in Table 3. It shows the performance measure for typical record 100, 117, and 119 and the average value for all 48 records.

Figure 4. Original, compressed, reconstructed and error ECG signal for AZTEC algorithm for typical record 119



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Performance measure for AZTEC on selected and averaged value for all 48 records from MIT-BIH database

Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	17.06	0	11.8	25.43	0.04	11.89	1.45
117	10.77	0	5.39	19.60	0.05	14.16	2
119	7.64	0.01	14.18	22.73	0.14	11.14	0.54
A*	10.65±3.66	0.01±0.01	20.02±19.38	28.31±20.77	0.09±0.09	12.19±4.10	0.83±0.73

A*- for all 48 records average± standard deviation

Advantage

- Higher compression ratio typically in order of 10.
- Applicable for automatic analysis such as QRS detection, having higher slope regions.

Limitations

- The quality of reconstructed signal is very low.
- The reconstructed signal is not acceptable to cardiologist.
- Needs random post processing to resemble the reconstructed signal to original signal.

Summary

The location of important diagnostic features are retained by location of bounds which basically determine the shape of ECG. The AZTEC algorithm keep only extrema (maxima and minima), plateau, and slope. During the reconstruction process plateau and slope are distinguished by introducing a flag of negative sign in time list.

CORTES

The CORTES algorithm (Abenstein & Tompkins, 1982) is the combination of TP and AZTEC method. Depending on the empirically determined threshold value, the length of the lines detected in AZTEC plateaus were retained otherwise the TP data point was saved. If length of the line is less than threshold than in that place of slope operation, TP method was employed. As TP method is efficient in high frequency region and same degree of distortion were added in both high and low frequency regions. Whereas distortion by TP method is tolerated in high frequency region compared to low frequency method. In low frequency region AZTEC algorithm provide low distortion by different length of plateau detection. The interpretation of length of the line is the related with the frequency of the signal, higher the length of the line lower the frequency leads smaller change in the signal and vice-versa. The brief procedure for *CORTES* is described in Algorithm 4.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 4. CORTES algorithm for ECG data compression

```
1:          INPUT:
2:          ecg is the original ECG signal
3:          OUTPUT:
4:          CORTES is the compressed ECG data
5:          Begin
6:          Detect plateau as in AZTEC ECG data compression Algorithm 3.
7:          Apply empirically determined threshold (TH) on the length of plateau
(L)
8:              If L < TH
9:                  TP method (Algorithm 1)
10:              Else
11:                  End
12:           End
```

Advantage

- Better than AZTEC and TP alone.

Limitations

- The quality of reconstructed signal is still low because of unavoidable distortion.

Summary

Cortes method is hybrid combination of TP and AZTEC method designed to preserve ECG features. The QRS feature of ECG is better preserved compared to AZTEC and isoelectric lines are compressed efficiently in comparison to TP method.

Fan

Fan algorithm for ECG data compression is based on finding the region bounded through the empirically determined threshold. The Fan algorithm draws lines between pairs of starting and ending points so that all intermediate samples are within some specified error tolerance, ε. The step-by-step basic Fan algorithm was illustrated by authors of (Jalaleddine et al., 1990; Tompkins, 1993). The brief procedure is presented in Algorithm 5. The visual representation for record number 119 in depicted in Figure 5. In Table 4 performance measure for Fan algorithm on selected and averaged value for all 48 records from MIT-BIH database is represented.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 5. Fan algorithm for ECG data compression

1: **INPUT:**
2: *ecg* is the original ECG signal
3: ε error tolerance
4: **OUTPUT:**
5: *Fan* is the compressed ECG data
6: **Begin**
7: X_0=*ecg(1);* Non-redundant permanent sample point or originating point
8: X_1= *ecg(1);* Taking second sample and drawing two lines (U_1, L_1)
U_1- passes through sample point (X_0, X_1+ ε)
L_1- passes through sample point (X_0, X_1- ε)
9: If the 3rd sample X_2 falls within the area bounded by two slopes, generate two new slopes (U_2, L_2) that passes through sample point (X_0, X_2+ ε) and (X_0, X_2- ε)
10: Compare the two pair of slopes and retain the most convergent pair. Which imply lower magnitude of U and higher of magnitude of L is retained.
11: Assign X_1= X_2 and read the next sample into X_2. X_2 always the most recent sample point and X_1 holds the sample immediately preceding to X_2.
12: Repeat the process by comparing X_2 to values of most convergent slopes.
13: If the sample point falls outside the area, save the line length '*len*' & final amplitude X_1, which then becomes the new originating point X_0 and repeat the process.
14: **End**

Matlab Simulation Program

Contents

- Initilization
- Fan algorithm for ECG data compression
- Reconstruction of ECG signal

```
clc;
clear all;
close all;
% Acquisition of ECG
load('100m.mat');
fs = 360;                              % Sampling Rate
time_sec=6*30;                        % time in number of seconds
ecg=(val(1,1:fs*time_sec))/200;  % Selecting lead II and removing the
base(1024) and gain(200)
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Initialization

```
j=1;i=1;k=1;
X0=ecg(i);                          %originating point
fan_compressed(j)=X0;               %save first permanent point
len=1;                              %initialize line length
Epsilon=.01;                          %error tolerance limit
l=length(ecg)-2;                    % length of ECG
T=1;
X1=ecg(i+1);                % Considering point 2
```

Fan Algorithm for ECG Data Compression

```
  XU1=X1+Epsilon;    XL1=X1-Epsilon;  % upper and lower bound of X1
 while (i<l-3)
     X2=ecg(i+2);           % get next sample point
%      XU1=X1+Epsilon;    XL1=X1-Epsilon;  % upper and lower bound of X1

     XU2 = (XU1 - X0)/T + XU1; % upper bound of X2
     XL2 = (XL1 - X0)/T + XL1; % lower bound of X2
     if (X2>=XL2 && X2<=XU2)   % within bound\

         % obtain the most restrictive bound
         if XU2 < X2 + Epsilon
             XU2=XU2;
         else
             XU2= X2 + Epsilon;
         end

         if XL2 > X2 - Epsilon
            XL2= XL2;
         else
             XL2=X2 - Epsilon;
         end

        len=len+1;
        X1 = X2 ; % X1 hold sample preceding X2
     else       % X2 out of bound, save line
       j=j+1; fan_compressed(j)=X1;  % save final amplitude
       k=k+1; fan_len(k)=len;        %  save line length

      % reset all variables
       X0=X1;
       X1=X2;
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
       len=1;
       T = 1 ;
XU1 = X1 + Epsilon ;
XL1 = X1 - Epsilon ;

     end
   i=i+len;

 end
```

Reconstruction of ECG Signal

```
z=1; %Sample number for reconstructed ECG
n=2; %To access the array storing lengths

for m=1:length(fan_compressed)-2
    len=fan_len(m+1);
    ecg1(z)=fan_compressed(m);
    for i=1:len
       ecg1(z+i)=ecg1(z); % reconstructed ECG
    end
   z=z+len;
end
```

Published with MATLAB® R2015a

Advantage

- Better than AZTEC and TP alone.
- Simple and faster implementation.
- Higher fidelity of reconstructed signal.

Limitations

- The quality of reconstructed signal is still low because of unavoidable distortion.

Summary

The fan algorithm incorporates the actual sample points which are bounded in the slopes. The three algorithms based on Scan-along polygonal approximation (SAPA) technique closely resemble to Fan algorithm.SAPA-2 proved best result among them. SAPA-2 calculates three slopes between origination sample and actual future sample point. The minor difference between SAPA-2 and Fan is the third slope calculation which is center slope criterion in place of actual sample value criteria.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Original, compressed, reconstructed and error ECG signal for Fan algorithm for typical record 119

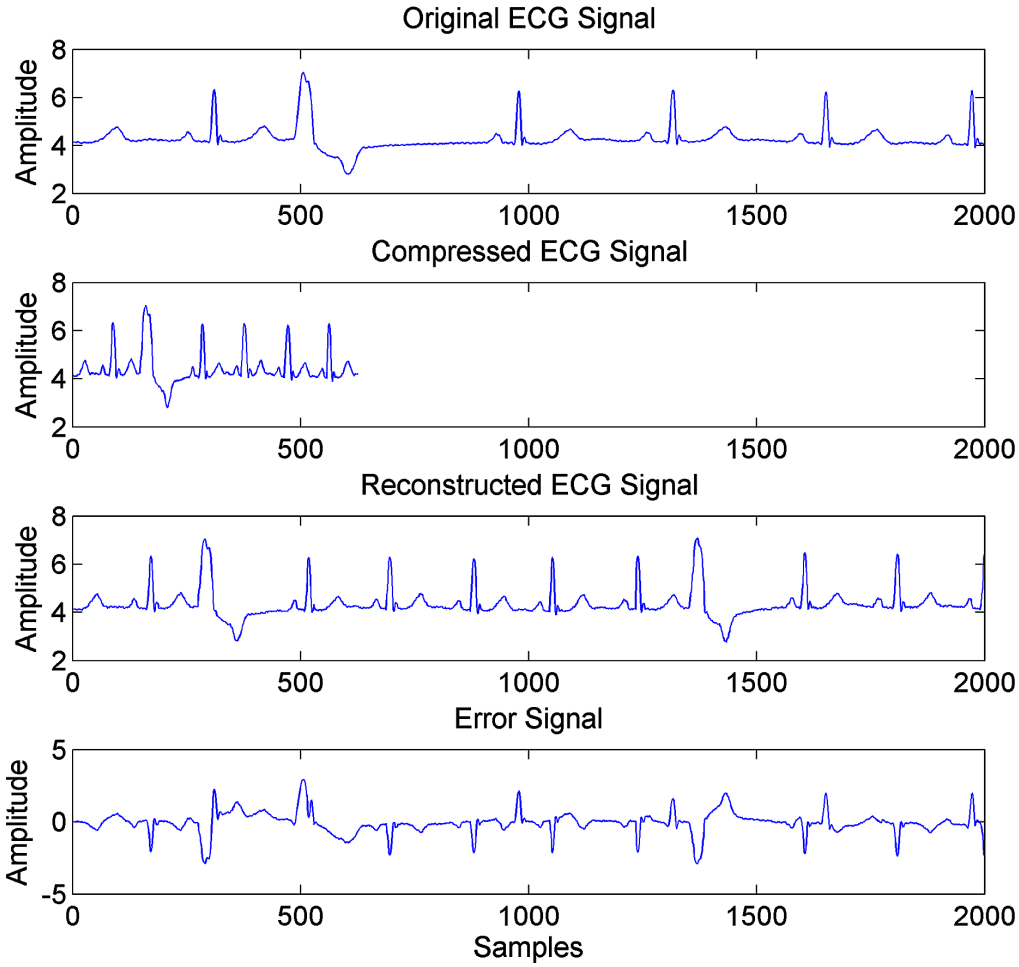


Table 4. Performance measure for Fan algorithm on selected and averaged value for all 48 records from MIT-BIH database

Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	2.23	0.03	6.63	180.87	0.32	-5.15	0.34
117	1.96	0.06	13.07	243.79	0.56	-7.74	0.15
119	1.96	0.08	20.01	164.58	0.86	-4.33	0.1
A*	1.9±0.11	0.06±±0.02	12.67±4.04	187.18±47.26	0.6±0.18	-5.22±1.88	0.17±0.05

A*- for all 48 records average± standard deviation

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Transformed Domain

The original ECG samples are subjected to a (linear) transformation in the new domain. The energy compaction property can be further utilized for compression. The basic property of energy compaction is “higher number of samples are clustered in certain region. In current scenario the certain region is near to the origin”. The data compression is achieved in any of domains are only through two ways 1) Thresholding and 2) Quantization.

1. **Thresholding:** There are different approaches for thresholding. When thresholding is applied in transformed domain only significant coefficient whose absolute value is above the threshold is retained. The selection of threshold may be user defined or via optimization algorithm.
2. **Quantization:** The thresholded significant coefficient is represented at lower bit rate through quantization scheme.

Walsh Transform

The Walsh transform amuses the ECG data compression application because it decomposes a signal into a set of orthogonal basis, rectangular waveforms of states are solely either +1 or -1. No multiplier is required and provides only real values as output. It is a asymmetrical and orthogonal transformation.

$$WAL(8) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{pmatrix}$$

The Walsh transform pair with forward and inverse for a signal x(n) of having length N are given by following

$$y_n = \frac{1}{N}\sum_{i=0}^{N-1} x_i WAL(n,i), n = 1,2.....,N-1 \tag{8}$$

$$x_i = \sum_{n=0}^{N-1} y_n WAL(n,i), n = 1,2.....,N-1 \tag{9}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As Walsh transform is symmetric and orthogonal the only difference between forward and its inverse is a multiplication of '1/N' the normalization factor.

The basic application of Walsh Transform in ECG data compression is previously demonstrated by authors of (Kulkarni, Kumar, & Verma, 1997) and its Matlab implementation is also available and explained in its application program. (http://in.mathworks.com/help/signal/examples/discrete-walsh-hadamard-transform.html). The visual representation for record number 119 in depicted in Figure 6. Table 5 and Table 6 represents the performance measure for Walsh transform on selected and averaged value for all 48 records from MIT-BIH database at Th=1 and 0.5 respectively.

Algorithm 6.

```
INPUT:
ecg is the original ECG signal
Define  Th (Threshold)
OUTPUT:
Walsh  transformed coefficient for representation of compressed ECG data
Begin
Take forward Walsh transform by using fwht command
Replace the Walsh coefficient with 0 whosoever is less than the empirically determined threshold
Reconstruct the original signal by using the inverse Walsh transform equation via iwht command in Matlab.
 Compute the performance measure
 End
```

Matlab Simulation Program

Contents

Initialization
Acquisition of ECG Signal
DCT Transform and thresholding
Reconstruction

```
clc
clear all
close all
```

Initialization

```
fs=360;                 % sampling frequency
sec=120;                % time duration of consideration
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
t=0:1/fs:sec-1/fs;       % relating time duration and sampling frequency with
number of samples
```

Acquisition of ECG Signal

```
load('117m.mat');  % load ecg data from database

ecg=val(1,1:fs*sec);  % fs*(number of seconds)                    % selecting
appropriate lead i.e. 1st row for lead II
mean_ecg=mean(ecg);
zxecg=ecg-mean(ecg);  % mean removal
```

DCT Transform and Thresholding

```
th=20*2;              % Emperically determined threshold
block_size=64*5;       % block size represents in how much size the DCT is
calculated over th signal

for i=1:length(zxecg)/block_size    % Block wise DCT coefficient computation
   dctzx((i*block_size)-(block_size-1):(i*block_size))=dct(zxecg((i*block_
size)-(block_size-1):i*block_size));
end

% Thresholding

ind_th=find(abs(dctzx)<th); % it calculates number of samples in Discrete Co-
sine Transformed version less the th threshold;
dctzx_th=dctzx;
dctzx_th(ind_th)=0;         % compressed DCT coefficients which is ready for
quantization and grouping
```

Reconstruction

```
for i=1:length(ecg)/block_size     % Inverse DCT computation
    inv_dctzx_th((i*block_size)-(block_size-1):(i*block_size))=idct(dctzx_
th((i*block_size)-(block_size-1):i*block_size));
end

est_ecg_th=mean_ecg+inv_dctzx_th; % Reconstructed ECG
```

Published with MATLAB® R2015a

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Original, compressed, reconstructed and error ECG signal for Walsh transform

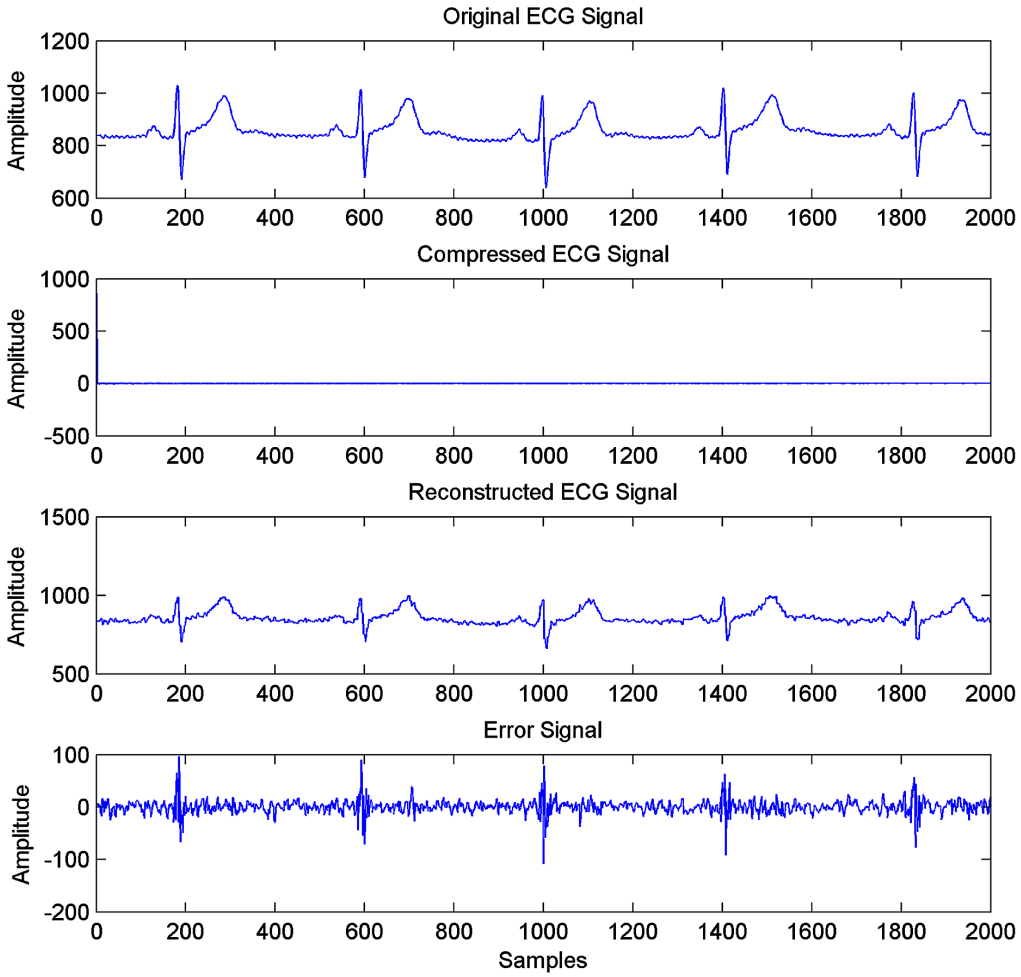


Table 5. Performance measure for Walsh transform on selected and averaged value for all 48 records from MIT-BIH database at Th=1

Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	28.44	30.71	2.31	65.56	22.15	3.67	12.34
117	21.79	11.87	2.3	42.93	19.88	7.34	9.47
119	9.79	12.38	2.72	24.62	23.52	12.17	3.6
A*	20.99±18.81	24.64±9.67	2.24±0.36	36.78±13.37	21.33±3.03	9.25±3.16	10.75±12.46

A*- for all 48 records average± standard deviation

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Performance measure for Walsh transformation selected and averaged value for all 48 records from MIT-BIH database at Th=0.5

Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	8.94	18.59	1.4	39.7	13.41	8.02	6.41
117	9.3	7.07	1.37	25.57	11.84	11.85	6.79
119	5.39	7.46	1.64	14.83	14.17	16.57	3.29
A*	8.67±4.82	15.2±6.29	1.37±0.18	22.75±9.13	13.01±1.47	13.52±3.37	6.84±4.88

Advantage

- Orthogonal transformation.
- Used power spectrum analysis, useful for characterizing non-linear signals & solving non-linear differential equations etc.

Limitations

- Computationally expensive with respect to wavelet transform.

Discrete Cosine Transform

The DCT is applied to ECG data compression field for improvement in the compression efficiency(Batista et al., 2001; Bendifallah, Benzid, & Boulemden, 2011). Researchers (Lee et al., 2011) have applied, the length of DCT block size should be same as the length of RR interval of ECG to capture the inter beat correlation. The basic formula for DCT is given below for computation of coefficients

$$y(k) = w(k)\sum_{n=1}^{L} X_s(n)\cos\left(\frac{\pi}{2L}(2n-1)(k-1)\right),\ k = 1,2,...,L$$

$$w(k) = \begin{cases} \dfrac{1}{\sqrt{L}}, & k = 1 \\ \sqrt{\dfrac{2}{L}}, & 2 \le k \le L \end{cases} \tag{10}$$

where $w(k)$ is scale factor

- After thresholding the significant DCT coefficient can be further quantized and grouped based on desired requirement.

The Algorithm 7 and its Matlab simulation code is basic implementation of DCT. After empirical estimation, a value equal to 40 was used for thresholding the raw DCT coefficients. Table 7 provide the values of performance measures for selected and average values of all records. The visual representation for record number is 119 in depicted in Figure 7.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 7. DCT coefficients computation for ECG data compression

```
INPUT:
ecg is the original ECG signal
Define block size and empirical threshold
OUTPUT:
dctzx_th is the DCT coefficients after thresholding
Begin
Remove average value in the ecg data points
Apply block wise DCT iteratively until last block
Finding the index in the DCT domain for coefficients to be greater
than threshold
 Retain the indexed DCT coefficient and replace the remaining sample
with zeros.
 End
```

Matlab Simulation Program

Contents

- Initialization
- Acquisition of ECG Signal
- DCT Transform and thresholding
- Reconstructtion

```
clc
clear all
close all
```

Initialization

```
fs=360;                 % sampling frequency
sec=120;                % time duration of consideration
t=0:1/fs:sec-1/fs;      % relating time duration and sampling frequency with
number of samples
```

Acquisition of ECG Signal

```
load('117m.mat');  % load ecg data from database
ecg=val(1,1:fs*sec);  % fs*(number of seconds)                  % selecting
appropriate lead i.e. 1st row for lead II
mean_ecg=mean(ecg);
zxecg=ecg-mean(ecg);  % mean removal
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DCT Transform and Thresholding

```
th=20*2;              % Empirically determined threshold
block_size=64*5;       % block size represents in how much size the DCT is
calculated over th signal

for i=1:length(zxecg)/block_size    % Block wise DCT coefficient computation
   dctzx((i*block_size)-(block_size-1):(i*block_size))=dct(zxecg((i*block_
size)-(block_size-1):i*block_size));
end

% Thresholding

ind_th=find(abs(dctzx)<th); % it calculates number of samples in Discrete Co-
sine Transformed version less the th threshold;
dctzx_th=dctzx;
dctzx_th(ind_th)=0;         % compressed DCT coefficients which is ready for
quantization and grouping
```

Reconstruction

```
for i=1:length(ecg)/block_size     % Inverse DCT computation
    inv_dctzx_th((i*block_size)-(block_size-1):(i*block_size))=idct(dctzx_
th((i*block_size)-(block_size-1):i*block_size));
end

est_ecg_th=mean_ecg+inv_dctzx_th; % Reconstructed ECG
```

Published with MATLAB® R2015a

Advantage

- The DCT is falling in the category of orthogonal transform.
- The DCT coefficient generally consists of large amount of near zero values which is main cause of compression.

Limitations

- While DCT operate on real valued input and typically real valued output. Thus there is a need of quantization step in block to achieve compression in same data format.
- Error propagation within the block size arises due to quantization and rounding during reconstruction process.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Original, compressed, reconstructed and error ECG signal for DCT

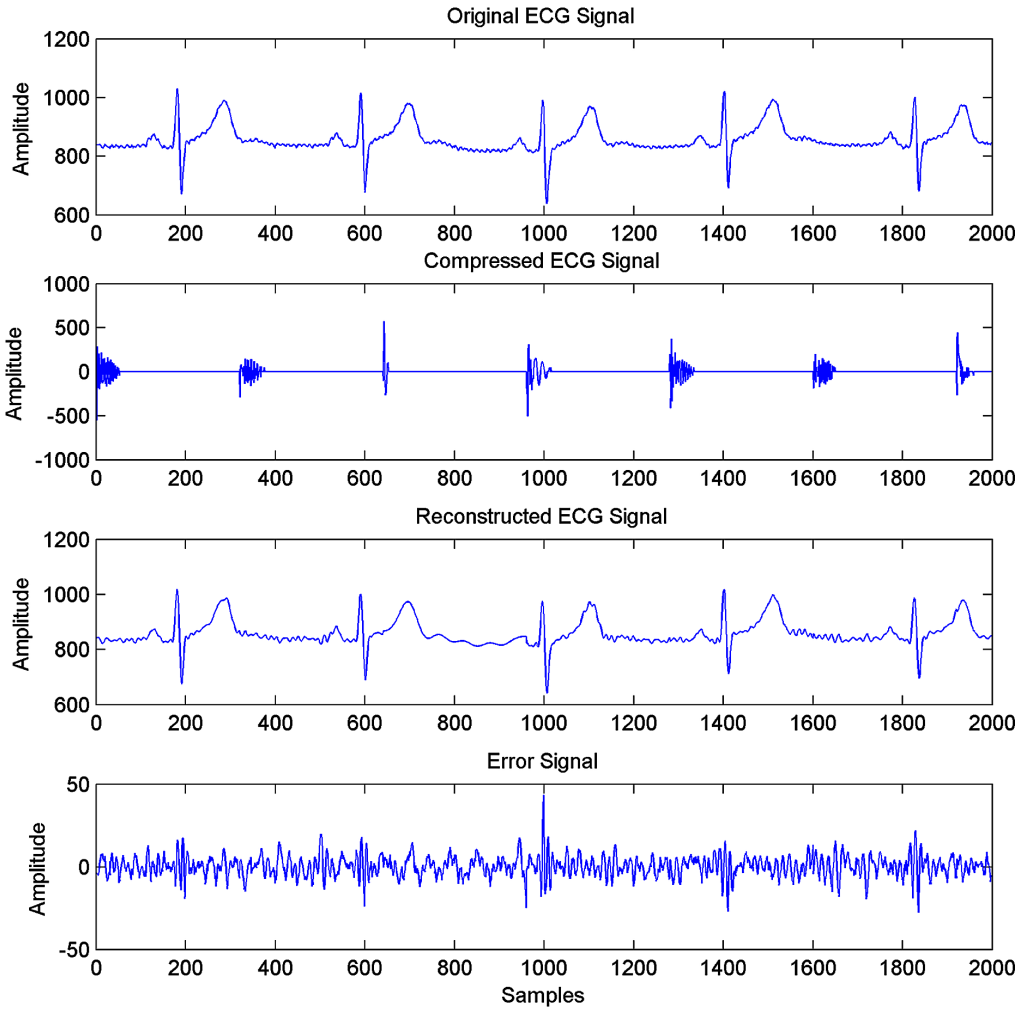


Table 7. Performance measure for CORTES on selected and averaged value for all 48 records from MIT-BIH database

Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	7.66	11.07	0.86	23.3	8.21	12.65	8.95
117	10.12	3.66	0.72	13.53	6.21	17.38	14.05
119	7.24	3.37	0.81	6.63	6.89	23.57	8.97
A*	9.30±3.19	8.90±4.61	0.78±0.11	12.82±6.53	7.43±1.14	18.74±3.76	12.31±4.95

A*- for all 48 records average± standard deviation

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Discrete Wavelet Transform

The Wavelet Transform is an effective and efficient tool for time-frequency localization. This time-frequency localization is also called as multi-resolution analysis (MRA) representation. The MRA is the basic fundamental approach of the wavelet transform that employed in wide variety of signal processing applications. The localized MRA approach has risen the processing capabilities in stationary and non-stationary domain. Further, it can be successfully applied to the signal feature detection, denoising, and compression. Several reference material for wavelet analysis are available (Daubechies, 1992; Stéphane Mallat, 2009). Wavelet transform is defined in both continuous as well in discrete domain.

The continuous wavelet transform (CWT) is one method of investigating the MRA of data whose frequency content varies with time. CWT maps *x(t)* which is a one dimensional signal to two-dimensional plane formulated by scale and translation defined as.

$$W(s,\tau) = \int x(t)\psi^{*}_{s,\tau}(t)dt \tag{11}$$

where *s* is scale and τ is measure of translation and $\psi^{*}_{s,\tau}(t)$ is the conjugate of mother wavelet formed by orthonormal basis functions. The CWT is computed for the particular scale (variable) by correlating the signal *x(t)* with translated mother wavelet $\psi^{*}_{s,\tau}(t)$. Higher the scale higher the frequency resolution or coarse frequency components and vice-versa. Its reverse is true for time resolution. The *x(t)* can be reconstructed back via inverse wavelet transform (Daubechies, 1992).

The explicit generation of mother wavelet from single wavelet ψ(*t*) is defined through scaling and translation

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{12}$$

The CWT applied on sampled data, with the smooth shifting process over the length of sampled data points, whereas the scaling is defined from the minimum (original signal scale) to a maximum as determined by the user, thus resulting in coarser resolutions in time and frequency domain. This improved resolution demands higher computational time and memory requirement for estimating the wavelet coefficients. The inherent redundancy and difficultly in analytical description of mother wavelet limits the usability of CWT in current scenario. Thus fast algorithm for wavelet coefficient computation is needed. To overcome above mentioned problem Discrete Wavelet Transform (DWT) has been profound. The DWT scales and translates only on discrete steps. In other words DWT is not time or frequency discrete rather than only translation and the scale step are discrete. Definitely it made a trade of between resolution and above mentioned problem but this compromise is widely acceptable.

The intuitively implementation for wavelet transform as the digital filter bank using Eq. 8 and 9 results in DWT defined as follows

$$DWT(m,k) = \frac{1}{\sqrt{s_0^m}}\sum_k x(k)\psi(s_0^{-m}n - k) \tag{13}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As mother wavelet takes various forms depending upon applications, a brief review of their utility for ECG signal processing is available in (Addison, 2005). Most wavelets can be formulated by estimating digital filter coefficients with its impulse response i.e. $\psi(s_0^{-m}n-k)$. The various diagrammatic illustration of forward and inverse DWT as filter banks is available in literature(Al-Shrouf, Abo-Zahhad, & Ahmed, 2003; Hilton, 1997; R. Kumar, Kumar, & Kumar, 2012) which all have the same meaning. The meaning is simple, which deals with the scaling function of the DWT represented as a tree of low and high pass filter bank, within each level only decimated (factor=2) low-pass signal will be transformed further. The subsequent larger scale decomposition results in lower resolution because of the lesser number of samples are available for analysis along with discarded high-frequency components analysis. A dyadic data size is a necessity of filter bank illustration for DWT. which means for *N* decomposition scale the data length must be 2^N. A wide range of literature for ECG data compression using wavelet transform is available with quantization and coding applied after decomposition (Al-Shrouf et al., 2003; Blanco-Velasco, Cruz-Roldán, Godino-Llorente, & Barner, 2007; Djohan, Nguyen, & Tompkins, 1995; Hilton, 1997; Isa, Noviyanto, Jatmiko, & Arymurthy, 2012; Wang & Meng, 2008).

In this Chapter, for ECG signal compression and reconstruction three procedure were adopted. Firstly *N,* level decomposition for given mother wavelet ('haar', 'db2', 'db10', & 'sym8') for input ECG signal was computed. Secondly, thresholding the detail coefficients at empirically or visually determined threshold value and lastly for reconstruction inverse wavelet transform is applied with the help of approximate and modified detail coefficient (exact low pass component and zeros replaced non- significant thresholded detail coefficient with same size) from all levels. The chapter focus is to implement DWT based ECG data compression so we left the quantization and coding part and demonstrated the wavelet based decomposition of ECG signal with its simulation program. The processing of ECG on DWT was operated for dyadic grid and *N=3* was selected for obtaining fixed *CR* by hard thresholding the detail coefficient. Table 8 represents the performance measure for DWT on selected and averaged value for all 48 records from MIT-BIH database over 'haar', 'db2', 'db10', & 'sym8' mother wavelet.

Algorithm 8.

```
INPUT:
ecg is the original ECG signal
OUTPUT:
Wavelet  transformed coefficient for representation of compressed
ECG data
Begin
Take forward wavelet transform by using dwt command
Replace the wavelet coefficient with 0 whosoever is being the part
of high frequency content or details component
Reconstruct the original signal by using the inverse wavelet trans-
form equation via idwt command in Matlab.
Compute the performance measure
 End
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Matlab Simulation Program

Contents

- Initialization
- Decomposition/ Compression part
- Synthesis/Reconstruction Part

```
clc
clear
close all
```

Initialization

```
fs=360;                  % sampling frequency
sec=60;                 % time duration of consideration
t=0:1/fs:sec-1/fs;       % relating time duration and sampling frequency with
number of samples
load('117m.mat');  % load ecg data from database
ecg=val(1,1:fs*sec);  % fs*(number of seconds)                    % selecting
appropriate lead i.e. 1st row for lead II
ecg_=(ecg(1:2^13)-1024)/200;
```

Decomposition/ Compression Part

```
mother_wave='sym8'; % Select any Mother Wavelet 'haar','db2', 'db10', or 'sym8'

[cA1 cD1]=dwt(ecg_,mother_wave);
[cA2 cD2]=dwt(cA1,mother_wave);
[cA3 cD3]=dwt(cA2,mother_wave);
[cA4 cD4]=dwt(cA3,mother_wave);
```

Synthesis/Reconstruction Part

```
X3 = idwt(cA4,zeros(1,length(cA4)),mother_wave);
X2 = idwt(X3,zeros(1,length(X3)),mother_wave);
X1 = idwt(X2,zeros(1,length(X2)),mother_wave);
X= idwt(X1,zeros(1,length(X1)),mother_wave);
ecg1=X; % Reconstructed ECG
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Performance measure for DWT on selected and averaged value for all 48 records from MIT-BIH database over 'haar', 'db2', 'db10', & 'sym8' mother wavelet

Haar							
Record	CR	PRD_1024	PRD	PRDN	RMS	SNR	QS
100	8	37.11	2.79	78.82	26.77	2.07	2.87
117	8	13.7	2.66	49.53	22.95	6.1	3.01
119	8	23.78	5.22	47.18	45.16	6.53	1.53
A*	8±0	36.74±13.11	3.67±1.56	53.4±13.17	34.77±14.02	5.74±2.32	2.69±1.41
db2							
100	8	35.15	2.64	74.64	25.35	2.54	3.03
117	8	12.62	2.45	45.63	21.14	6.81	3.27
119	8	21.09	4.63	41.85	40.05	7.57	1.73
A*	8±0	32.86±12.97	3.26±1.48	47.96±14.47	30.86±13.13	6.82±2.87	3.04±1.55
db10							
100	8	35.8	2.69	76.04	25.83	2.38	2.98
117	8	12.47	2.42	45.11	20.9	6.92	3.31
119	8	19.87	4.36	39.44	37.75	8.08	1.83
A*	8±0	30.72±13.29	3.03±1.43	45.09±15.86	28.65±12.79	7.55±3.48	3.34±1.8
sym8							
100	8	35.29	2.65	74.98	25.45	2.5	3.02
117	8	12.78	2.48	46.26	21.42	6.7	3.23
119	8	19.49	4.28	38.7	37.02	8.25	1.87
A*	8±0	30.87±13.61	3.04±1.45	45.27±16.24	28.74±12.99	7.55±3.58	3.37±1.92

A*- for all 48 records average± standard deviation

Advantage

- Ability to handle non-stationary data.
- Good localization both in time and frequency domain.
- Better ECG performance measure.
- Ability of automatic removal of outliers.

Limitations

- Functional form representation of many wavelets is difficult.
- The DWT is shift sensitive because input-signal shifts generate unpredictable changes in DWT coefficients.
- The DWT suffers from poor directionality because DWT coefficients reveal only three spatial orientations.
- The DWT analysis lacks the phase information that accurately describes non-stationary signal behavior.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Summary

The formalism of the one-dimensional discrete wavelet transform (DWT) based on different mother wavelet is outlined for ECG data decomposition.

Parameter Extraction Method

The parameter extraction method uses pre-processing tool to obtain a set of set-points or features (typical cycles, extreme locations, etc.) that applied to model the original signal. These features are further quantized using the scalar or the vector quantizer and then some lossless coder is utilized to increase the compression efficiency. Reconstruction of original data is carried out by adopting suitable interpolation scheme. The parameter extraction technique is based on the assumption that the ECG signal is formulated through replication of the number of beats at the frequent rate. Review of some ECG data compression method can be readily found in reference (Jalaleddine et al., 1990; Sabarimalai Sur & Dandapat, 2014). Parameter extraction ECG compression methods mainly classified into prediction models(Al-Shrouf et al., 2003; Arnavut, 2007; Cohen & Zigel, 1998; Nave & Cohen, 1993; Ramakrishnan, 1997). Others are also significant for ECG compression includes scalar/ vector quantization (Cohen & Zigel, 1998; Huang, Wang, & Chen, 2013; Shaou-Gang & Yen, 2001) and template matching method (Paggetti et al., 1994).

In Linear prediction model method the instantaneous sample ($\hat{x}$ (n)) is forecasted by (a) weighted combination past (p) samples.

$$\hat{x}(n) = -a(2)x(n-1) - a(3)x(n-2) - ... - a(p+1)x(n-p) \tag{14}$$

Then, the error among the forecasted and original samples is sent or stored in place of the original samples. The error is quantized using any efficient quantizer (scalar/vector) which is further encoded with lossless encoding. In this process major compression is achieved in two stages first in by reducing the dynamic range and secondly in process of quantization.

The vector quantization is mainly based on the code book creation and transmitting the only codebook number instead of the wide range of samples which are assigned to that particular codebook number. Codebook may be the signal dependent or universal which is available to both of encoder and decoder.

The template matching based parameter extraction method for ECG compression is similar to vector quantization. The template matching for ECG data compression include steps of the beat segmentation, normalization, matching and subtraction with the coding of the residual signal.

The MATLAB code can be made by implementing the any of individual diverse methodology available in the literature.

CONCLUSION

In this chapter, the implementation of different selected techniques for ECG data compression in MATLAB was presented. The implemented ECG compression algorithm may vary with the exact algorithm for particular technique available in the literature due to the focus of this chapter is on the basic methodology of technique rather than pre-or-post algorithm involved with methods. The selected techniques

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

from direct ECG compression methods are TP, AZTEC, Fan, and Cortes. Moreover selected techniques from transformed ECG compression methods are Walsh Transform, DCT, and Wavelet transform. The performance measure for implemented algorithm is also represented. This study seems to be of selection for the appropriate algorithm as needed.

REFERENCES

Abenstein, J. P., & Tompkins, W. J. (1982). A new data-reduction algorithm for real-time ECG analysis. *IEEE Transactions on Bio-Medical Engineering*, *29*(1), 43–48. doi:10.1109/TBME.1982.324962 PMID:7076268

Addison, P. S. (2005). Wavelet transforms and the ECG: A review. *Physiological Measurement*, *26*(5), 155–199. doi:10.1088/0967-3334/26/5/R01 PMID:16088052

Al-Shrouf, A., Abo-Zahhad, M., & Ahmed, S. M. (2003). A novel compression algorithm for electrocardiogram signals based on the linear prediction of the wavelet coefficients. *Digital Signal Processing*, *13*(4), 604–622. doi:10.1016/S1051-2004(02)00031-3

Arnavut, Z. (2007). ECG signal compression based on Burrows-Wheeler transformation and inversion ranks of linear prediction. *IEEE Transactions on Bio-Medical Engineering*, *54*(3), 410–418. doi:10.1109/TBME.2006.888820 PMID:17355052

Baibars, M., Kanjwal, K., & Marine, J. E. (2014). AF Detected on Implanted Cardiac Implantable Electronic Devices: Is There a Threshold for Thromboembolic Risk? *Current Treatment Options in Cardiovascular Medicine*, *16*(3), 289. doi:10.100711936-013-0289-7 PMID:24500679

Batista, L. V., Melcher, E. U. K., & Carvalho, L. C. (2001). Compression of ECG signals by optimized quantization of discrete cosine transform coefficients. *Medical Engineering & Physics*, *23*(2), 127–134. doi:10.1016/S1350-4533(01)00030-3 PMID:11413065

Bendifallah, A., Benzid, R., & Boulemden, M. (2011). Improved ECG compression method using discrete cosine transform. *Electronics Letters*, *47*(2), 87. doi:10.1049/el.2010.3191

Blanco-Velasco, M., Cruz-Roldán, F., Godino-Llorente, J. I., & Barner, K. E. (2007). Wavelet packets feasibility study for the design of an ECG compressor. *IEEE Transactions on Bio-Medical Engineering*, *54*(4), 766–769. doi:10.1109/TBME.2006.889176 PMID:17405386

Blanco-Velasco, M., Cruz-Roldán, F., Godino-Llorente, J. I., Blanco-Velasco, J., Armiens-Aparicio, C., & López-Ferreras, F. (2005). On the use of PRD and CR parameters for ECG compression. *Medical Engineering & Physics*, *27*(9), 798–802. doi:10.1016/j.medengphy.2005.02.007 PMID:15869896

Cheng, A. C. (2015). *Real-Time Cardiovascular Diseases Detection on a Smartphone*. Academic Press.

Cohen, A., & Zigel, Y. (1998). Compression of multichannel ECG through multichannel long-term prediction. *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine & Biology Society, 17*(1), 109–115. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9460626

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cox, J. R., Nolle, F. M., Fozzard, H. A., & Oliver, G. C. (1968). AZTEC, a preprocessing program for real-time ECG rhythm analysis. *IEEE Transactions on Bio-Medical Engineering*, *15*(April), 128–129. doi:10.1109/TBME.1968.4502549 PMID:5648088

Daubechies, I. (1992). Ten Lectures on Wavelets. *IEEE Symposium on Computer Based Medical Systems* (Vol. 61).

Djohan, A., Nguyen, T. Q., & Tompkins, W. J. (1995). ECG compression using discrete symmetric wavelet transform. In *IEEE International Conference on Engineering in Medicine and Biology Society* (pp. 167–168). Montreal, Canada: IEEE. 10.1109/IEMBS.1995.575053

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, *101*(23), e215–e220. doi:10.1161/01.CIR.101.23.e215 PMID:10851218

Hebert, M., Korabek, B., & Scott, R. E. (2006). Moving research into practice: A decision framework for integrating home telehealth into chronic illness care. *International Journal of Medical Informatics*, *75*(12), 786–794. doi:10.1016/j.ijmedinf.2006.05.041 PMID:16872892

Hilton, M. L. (1997). Wavelet and wavelet packet compression of electrocardiograms. *IEEE Transactions on Bio-Medical Engineering*, *44*(5), 394–402. doi:10.1109/10.568915 PMID:9125824

Horspool, R. N., & Windels, W. J. (1994). An LZ approach to ECG compression. In *Computer-Based Medical Systems, 1994., Proceedings 1994 IEEE Seventh Symposium on* (pp. 71–76). http://doi.org/10.1109/CBMS.1994.315990

Huang, B., Wang, Y., & Chen, J. (2013). ECG compression using the context modeling arithmetic coding with dynamic learning vector–scalar quantization. *Biomedical Signal Processing and Control*, *8*(1), 59–65. doi:10.1016/j.bspc.2012.04.003

Isa, S. M., Noviyanto, A., Jatmiko, W., & Arymurthy, A. M. (2012). The Effect of Electrocardiogram Signal Compression Using Beat Reordering and SPIHT on Automatic Sleep Stage Classification. *Procedia Engineering, 41*, 888–896. http://doi.org/ doi:10.1016/j.proeng.2012.07.259

Jalaleddine, S. M., Hutchens, C. G., Strattan, R. D., & Coberly, W. A. (1990). ECG data compression techniques-a unified approach. *IEEE Transactions on Bio-Medical Engineering*, *37*(4), 329–343. doi:10.1109/10.52340 PMID:2186997

Koski, A. (1997). Lossless ECG encoding. *Computer Methods and Programs in Biomedicine*, *52*(1), 23–33. doi:10.1016/S0169-2607(96)01779-8 PMID:9034667

Kulkarni, P. K., Kumar, V., & Verma, H. K. (1997). ECG data compression using fast Walsh transform and its clinical acceptability. *International Journal of Systems Science*, *28*(8), 831–836. doi:10.1080/00207729708929443

Kumar, R., Kumar, A., & Kumar, R. (2012). Electrocardiogram Signal Compression Using Beta Wavelets. *Journal of Mathematical Modelling and Algorithms*, *11*(3), 235–248. doi:10.100710852-012-9181-9

Kumar, V., Saxena, S. C., & Giri, V. K. (2006). Direct data compression of ECG signal for telemedicine. *International Journal of Systems Science*, *37*(1), 45–63. doi:10.1080/00319100500412337

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lee, S., Kim, J., & Lee, M. (2011). A real-time ECG data compression and transmission algorithm for an e-health device. *IEEE Transactions on Bio-Medical Engineering*, *58*(9), 2448–2455. doi:10.1109/TBME.2011.2156794 PMID:21606020

Ma, J., Zhang, T., & Dong, M. (2015). A Novel ECG data compression method using adaptive fourier decomposition with security guarantee in e-Health applications. *IEEE Journal of Biomedical and Health Informatics*, *19*(3), 986–994. PMID:25222961

Mallat, S. (2009). A wavelet tour of signal processing (S. Mallat, Ed.). Academic Press.

Mueller, W. C. (1977). Arrhythmia detection program for an ambulatory ECG monitor. *Biomedical Sciences Instrumentation*, *14*(1), 81–85. PMID:687739

Nave, G., & Cohen, A. (1993). ECG compression using long-term prediction. *IEEE Transactions on Bio-Medical Engineering*, *40*(9), 877–885. doi:10.1109/10.245608 PMID:8288278

Paggetti, C., Lusini, M., Varanini, M., Taddei, A., Marchesi, C., & Sistemi, D. (1994). *A Multichannel Template Based Data Compression Algorithm*. Academic Press.

Ramakrishnan, A. G., & Saha, S. (1997). ECG Coding by Wavelet-Based Linear Prediction. *IEEE Transactions on Bio-Medical Engineering*, *44*(12), 1253–1261. doi:10.1109/10.649997 PMID:9401225

Sabarimalai Sur, M., & Dandapat, S. (2014). Wavelet-based electrocardiogram signal compression methods and their performances: A prospective review. *Biomedical Signal Processing and Control*, *14*(1), 73–107. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1746809414000998

Sayood, K. (2010). Introduction to data compression. (K. Sayood, Ed.). Elsevier Inc.

Shaou-Gang, M., & Yen, H.-L. (2001). Multichannel ECG compression using multichannel adaptive vector quantization. *IEEE Transactions on Bio-Medical Engineering*, *48*(10), 1203–1207. doi:10.1109/10.951524 PMID:11585045

Singh, B., & Singh, D. (2012). Effect of threshold value r on multiscale entropy based heart rate variability. *Cardiovascular Engineering and Technology*, *3*(2), 211–216. doi:10.100713239-012-0082-x

Tompkins, W. J. (1993). *Biomedical digital signal processing*. Editorial Prentice Hall.

Wang, X., & Meng, J. (2008). A 2-D ECG compression algorithm based on wavelet transform and vector quantization. *Digital Signal Processing*, *18*(2), 179–188. doi:10.1016/j.dsp.2007.03.003

This research was previously published in Computational Tools and Techniques for Biomedical Signal Processing edited by Butta Singh; pages 272-313; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 50
A Study on Lifetime Enhancement and Reliability in Wearable Wireless Body Area Networks

Saranya Vasanthamani
Sri Krishna College of Engineering and Technology, India

ABSTRACT

The wireless body area network (WBAN) which consists of wearable or implantable sensor nodes, is a technology that enables pervasive observing and delivery of health related information and services. The radio-enabled implantable medical devices offer a revolutionary set of applications among which we can point to precision drug distribution, smart endoscope capsules, glucose level observers and eye pressure detecting systems. Devices with WBAN are generally battery powered due to sensitivity and criticality of the data carried and handled by WBAN, reliability becomes a critical issues. WBAN loads a high degree of reliability as it openly affects the quality of patient observing. Undetected life-threatening circumstances can lead to death. A main requirement is that the health care professionals receive the monitored data correctly in emergency situations. The major objective is to achieve a reliable network with minimum delay and maximum throughput while considering power consumption by reducing unnecessary communication.

1. INTRODUCTION

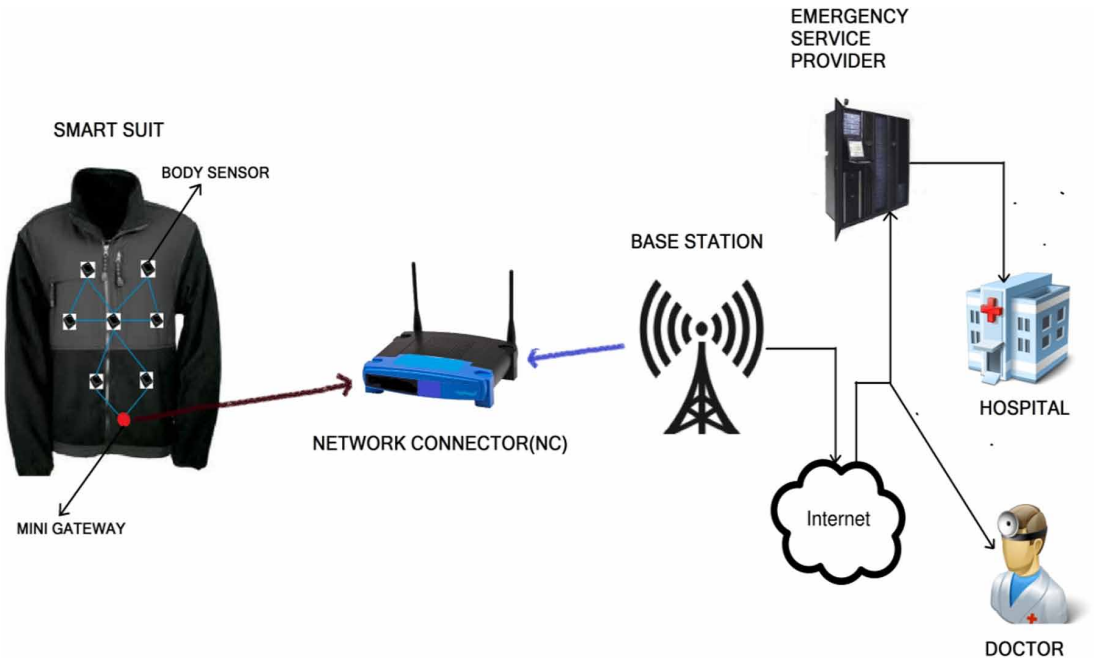
Wireless technology has advanced to become a vital part of our lives starting from mobile communication to health care departments. Recently, there has been growing interest from system inventors application and researchers on a newly designed type of network architecture usually known as body sensor networks (BSNs) or body area networks (BANs), one made feasible by novel advances on ultra-low-power, lightweight, small-size and intelligent observing wearable sensors. In BANs, sensors are

DOI: 10.4018/978-1-7998-1204-3.ch050

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Data transmission in WBSN



used to constantly monitor human's functional activities and movements, such as health position and movement pattern. A wireless healthcare application offers and brings many benefits and challenges to the health care sector. These benefits provide a convenient-environment that can monitor the daily lives and medical situations of patients at anytime, anywhere and without limitations. The Figure 1 shows how WBAN provides a wireless connection between these devices and a Personal Digital Assistant (PDA) or a smartphone, which is responsible for the connection with other networks.For example, the obtained data can be forwarded to a hospital server.

1.1 Wireless BSN Scenario

Wireless BSN applications are in great demand in medical care (Movassaghi, Abolhasan, Lipman, Smith, & Jamalipour, 2014; Cavallari, Martelli, Rosini, Buratti, and Verdone, 2014; Davies, Sanjay, & Mohana, 2014), sports and entertainment (Conroy et al., 2009), the military-industrial sector (Kavitha, Balapriya, & Sundrarajan, 2016), and the social public field (Ahmad Dogar & Zafar, 2012; Akyildiz, Su, Sankarasubramaniam, & Cayirci, 2002; Patel & Wang, 2010), and BSNs have gradually become a research hotspot. BSNs is a type of WSN which is formed by physiological factors of sensors placed in the human body or on the body surface or around the body. The key performances it covers are sensors, data fusion, and network communication. It focuses on the advancements in universal health care, disease monitoring, and prevention solution, but also an essential component of the so-called Internet of Things. Its fore most purpose is to make available an integrated ubiquitous computing hardware, software, and wireless communication technology platform, and a vital situation for the imminent improvement of ubiquitous healthcare observing systems (Khan & Yuce, 2010). BSNs initiated from WSNs, so there are many resemblances between them. However, the features are consistently different as of their different

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

application resolutions. Initially, allowing for network deployment, WSNs can be deployed to the unreachable environments, such as forests, swamps or mountains. Several redundant nodes are positioned in the positions indicated above to resolve the problem of node failures, so node density is greater,however, BSN nodes are positioned in, on or around the human body, so the total number of nodes is usually up to a few dozens. Each node confirms the correctness of observing consequences by its robustness (Crosby, Ghosh, Murimi, Chin, & Craig, 2012). Moreover, considering attributes, nodes in WSNs accomplish the similar functions and have the identical properties. The size of nodes is not very critical. Formerly the node is deployed, it will probably no longer need to be relocated. Agreeing to the dissimilar physical signals collected, BSN applications ensure different sensor types (Khan, Hussain, & Kwak, 2009). Furthermore, the necessities of BSN node design are comparatively high. The node dimensions must be small enough, and the nodes must have high wearability and high biocompatibility (Carrano, Passos, Magalhaes, & Albuquerque, 2013). Due to the locations, the nodes are deployed, they will move as the human body moves. Furthermore, considering energy supply, WSNs and BSNs can be battery-powered. The earlier, positioned outdoors, can also be powered by wind energy or solar energy, while the latter can also be powered by kinetic energy and heat (Reddy, Reddy, & Reddy, 2013). Lastly, allowing for data transmission, the transfer rates of WSNs are almost the same, but those of BSNs are different, as the data type and channel task are different among nodes on the body surface and in the BSNs (Khan, Hussain, & Kwak, 2009). Additionally, BSN deployed in the human body are for observing human physiological data, which are subject to user's personal safety and privacy protection issues. Therefore, QoS and the real-time prosperity of data transmission must be considered (Ben Elhadj, Chaari Fourati, & Kamoun, 2012; FCC, 2013; Seyedi, Kibret, Lai, & Faulkner, 2013).

1.1.1 Implanted Node

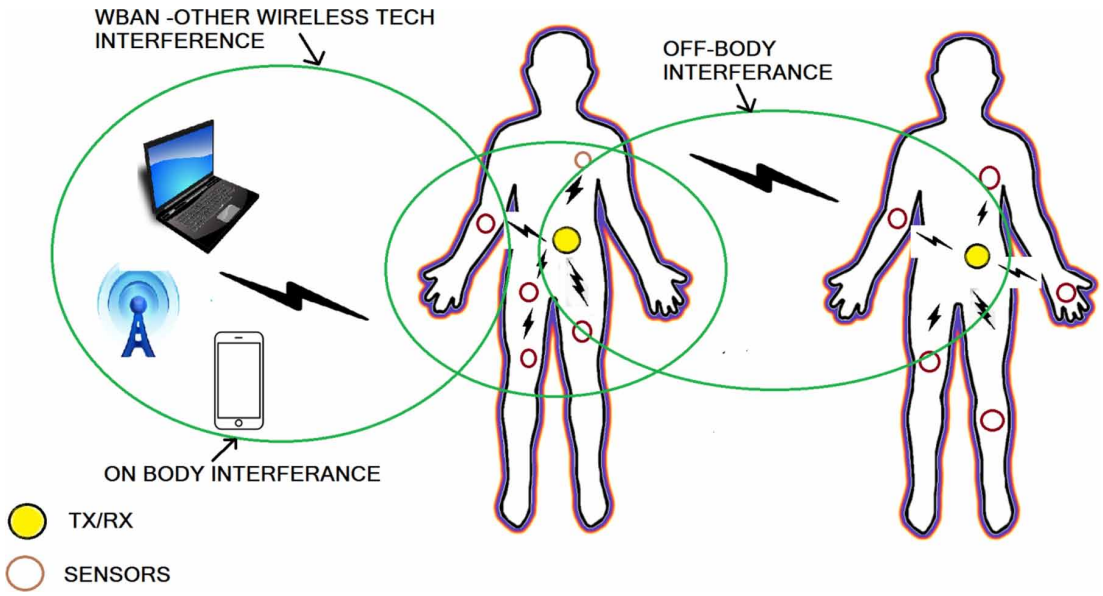
Implanted nodes are used to analyze the health factors with the biosensors which in need to be in the close touching base with the skin, and every now and then even inside the human body. Implantable biosensors are an important class of biosensors constructed on their capability to continuously measure metabolite levels, without the prerequisite for patient intervention and irrespective of the patient's physiological state (sleep, rest, etc.) For example, implantable biosensors symbolize an extremely necessary proposition for diabetes management which presently relies on data obtained by using test strips blood from finger pricking, a practice which is not only painful, but as well is incapable of reflecting the global direction, trends, and patterns related with daily habits.

A startup of a broad research effort predictable at developing implantable biosensors for continuous observing of multiple biologically relevant metabolites, but not wearable sensors.Additional classes of implantable devices which ensure be situated intensively investigated comprise of sensors for nerve stimulation can ease acute pain sensors to the detecting electric signals in brain and sensors to monitor biological analysis in the brain with implanted drug delivery systems for controlled delivery at the site of pain and stress (Tang, Tummala, Gupta, & Schwiebert, 2005).

In the course of the end of the last era, there has been a major increase in the number of various wearable health observing devices, extending from simple pulse monitors, activity monitors, and portable Holter monitors, to sophisticated and exclusive implantable sensors. IWBANs (implantable wireless body area networks) are more desirable than WWBAN (wearable wireless body area network) for numerous advantages. WWBANs have some disadvantages: they limit the mobility of the patients; in accumulation, they can cause skin infections, thus donating poor health conditions. Regardless of the wireless connec-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Representations of implanted sensors



tion is not an essential requirement for observing physiological parameters from implanted sensors. This problem is measured as one of the main incentives for the trend in modern biomedical implanted systems to use wireless technology (Tang, Tummala, Gupta, & Schwiebert, 2005). In Binkley (2003) explored a vision of the near future when one single device will be able to create a WSN with a large number of nodes, which are put on inside and outside the body may be one or the other predetermined or aimlessly, in accord with the application. This visualization can only be achieved over a broad communication standard for wireless telemetry link. Typical hardware and software architecture can support compatible devices, which are predictable to meaningfully affect the succeeding generation of healthcare systems. Certain of these devices can then be incorporated into the wireless body area network, providing novel opportunities for technology to monitor the health status.

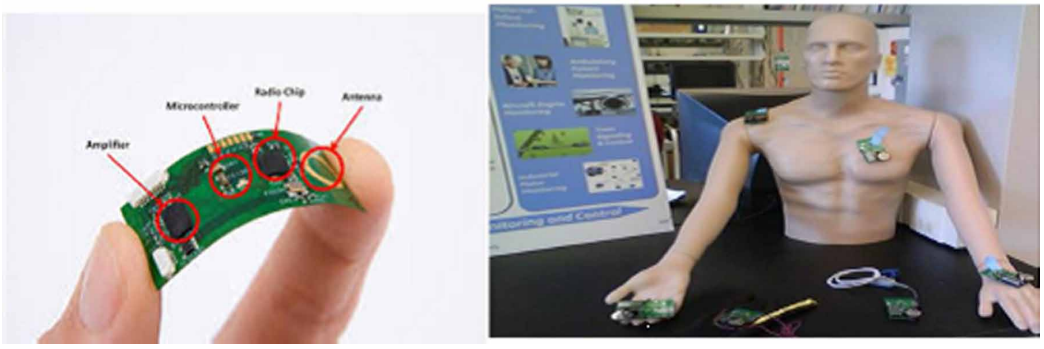
1.1.2 Non Implanted Node

WSN technology offers a wide range of welfares to patients, medical staff, and society through uninterrupted observing in an ambulatory setting, premature detection of abnormal circumstances, vision rehabilitation, and potential discovery of knowledge through data mining of all assembled information. This part shows exactly how to use the planning and mechanism of WWBANs as well as main arrangement enabling unobtrusive, continuous, daily observing of health. In accumulation, this section defines some important implementation issues.

Wearable health observing systems agree the individual follow closely the changes in her or his vital functions and provide feedback for maintaining optimal health status. If integrated into the telemedicine system, such systems can alert medical personnel when serious changes occur. In accumulation with medical centers, patients could be benefited for uninterrupted long-term observing as a portion of a diagnostic practice. We can accomplish optimal maintenance of a chronic condition or can be monitored during the recovery period after the critical event or surgical procedure. Durable health observing can

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Non-implanted sensors nodes



capture the diurnal and circadian variations in physiological signals. These changes, for example, are a very good indicator of cardiac recovery of patients after myocardial infarction (Otto, Milenković, Sanders, & Jovanov, 2006). Long-term observing can also confirm adherence to treatment guidelines or help monitor the effects of drug therapy. Other patients may also benefit from these systems; for example, monitors can be used during physical restoration after hip or knee surgeries, stroke restoration, or brain trauma restoration. The Internet of Things (IoT) is gaining a speedily growing attention in many disciplines, in particular for personalised healthcare. In the meantime, body area sensor network (BASN) under the IoT structure has been widely functional for ubiquitous health monitoring (Wu, Wu, Redouté, & Yuce, 2017; Negra, Jemili, & Belghith, 2016).

The practice of wearable sensors for observing several health-related biometric parameters in daily activities is attracting more interest in recent times. Several individuals are familiar with the use of devices such as wearable heart rate monitors and pedometers for medical reasons or as part of a fitness administration. The architecture includes the major stage to cover the number of wireless nodes of medical sensor that are integrated into WWBANs. Respectively, each sensor node sense, sample, and process one or more physiological signals. For example, an electro cardiogram sensor can be used to monitor heart action, an electromyogram sensor for observing muscle activity, an electro encephalogram sensor for observing brain electrical activity, a blood pressure sensor for observing blood pressure, a tilt sensor for observing trunk situation, and a breathing sensor for observing breathing; and motion sensors can be used to categorize the user's status and to estimate her or his level of activity.

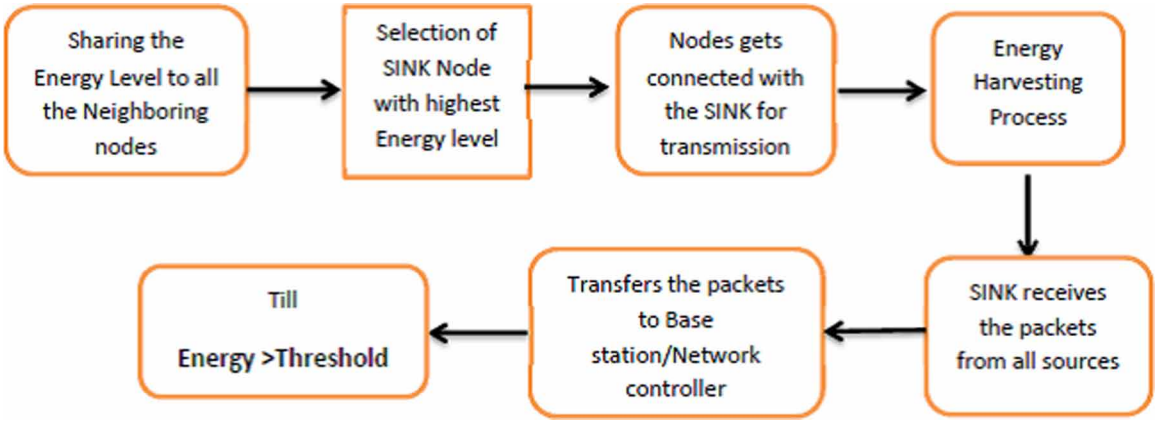
1.2 Topology of the Node Deployment

The deployment of nodes for the data transmission in a network is based on the Energy level of the nodes and the node with the maximum energy is considered as a SINK node.

Initially, the energy level of all the neighboring nodes is analyzed for the selection of a SINK node which holds the highest energy level. After the selection of the SINK, all the other nodes get connected with the SINK till the energy level reaches the minimum threshold level. The packet transmission takes place with the assistance of the SINK acting as an intermediate with the network controller. The same process is repeated till the SINK energy level decreases. Even after reaching the minimum energy level the transmission has to be continued then reelection of another SINK node will to processed for a long-term communication with the same procedure to be continued.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Data communication using SINK



2. LIFETIME ENHANCEMENT METHODS

2.1 Concept of Network Lifetime

In WBANs, energy is consumed mainly on transmission, sensing and unusual energy unused including idle listening, collision, and overhearing. Wireless Body Area Network (WBAN) is a collection of low-power, miniaturized, invasive/non-invasive lightweight wireless sensor nodes that monitor the human body functions and the surrounding environment

2.2 Network Model

WBAN devices are implanted inside the body or might be surface-mounted on the body in a fixed position or else may be accompanied devices which humans can carry in dissimilar positions, in dress pockets, by hand or in bags.Whereas there is a development of tiny devices, in specific, networks consisting of several tiny body sensor units (BSUs) organized with a particular body central unit (BCU).The observing is done in both the places at home and office with these nodes. The development in WBAN technology implements communications on, near, and around the human body and they can transfer information through gateways to reach longer ranges. Through gateway devices, it is possible to connect the wearable devices on the human body to the internet. This way, medical professionals can access patient data online using the internet independent of the patient location.

2.3 Energy Consumption Model

Medical information collected by sensors on the patient's body (WPAN-Wireless Personal Area Networks) is displayed on a bedside monitor. This information is also transmitted to another hospital location for remote monitoring, e.g., a nurses' station) In case of emergency, when the patient is moved from his/her room to the intensive care unit, these communications need to be maintained.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Network model

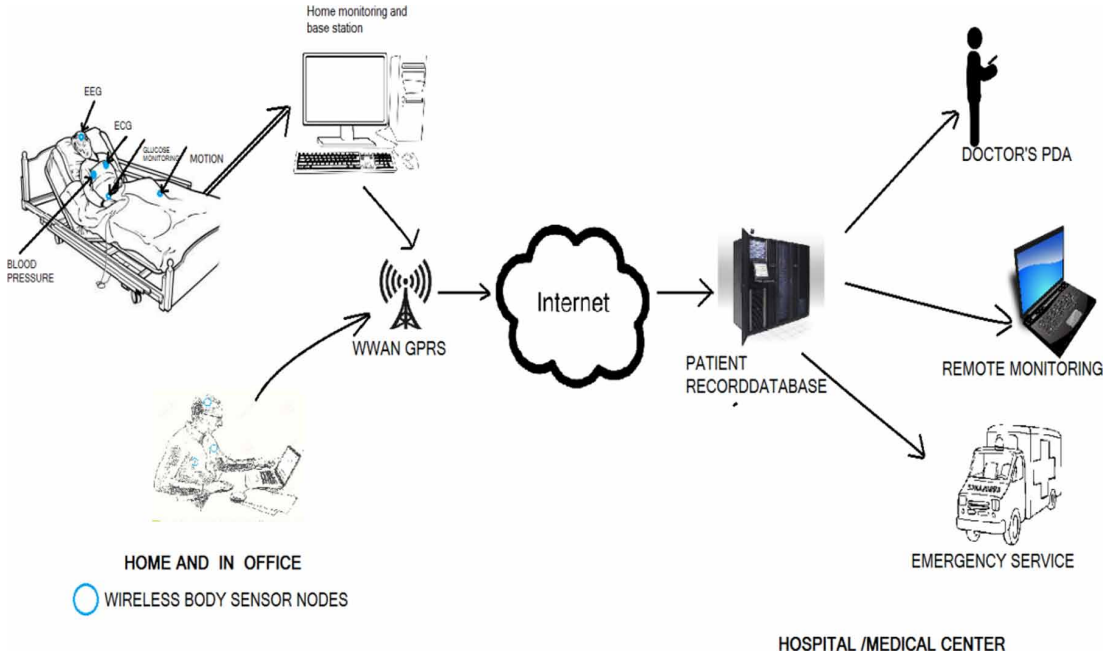
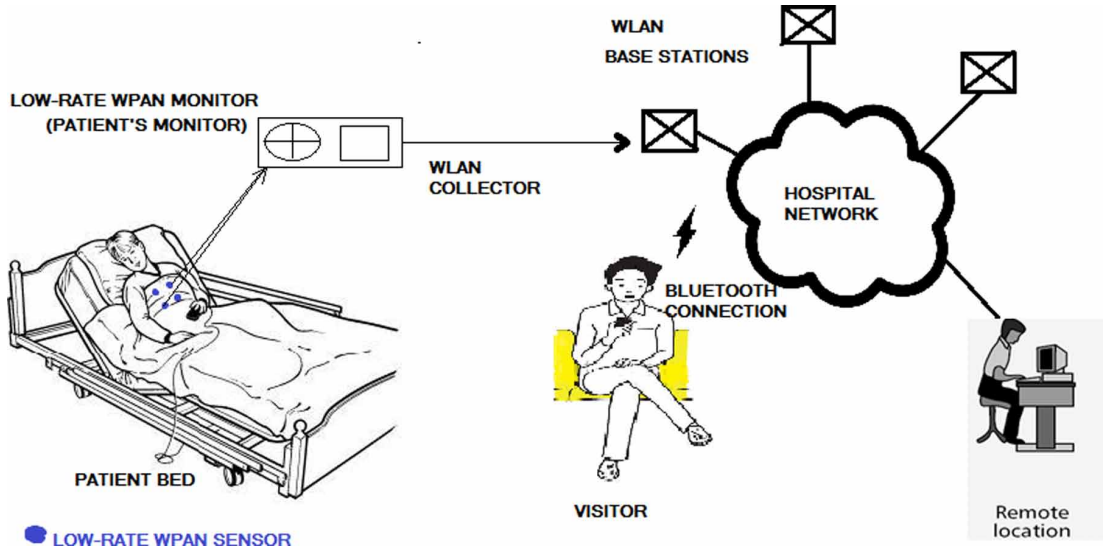


Figure 6. Energy consumption and node deployment using WPAN



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$d_{i,j}$: The distance between each pair of sensor nodes and the distance between sensor nodes and the controller.
p_t: The transmission path state (LOS/NLOS-Line of Sight/Non Line of Sight) between each pair of nodes including the controller.
ETX_{elec}, ERX_{elec}, E_{amp}: The energy consumption parameters.
P_L: The packet length in the network.
E_s: The energy storage condition of each sensor node in the network.

The energy cost to guarantee reliable transmission during transmitting and receiving are described in Equations (1) and (2) as follows:

$$Etx_{(PL,d,pt)} = ETX_{elec} * P_L + E_{amp} * P_L * d_{pt} \tag{1}$$

$$Erx_{(k)} = ERX_{elec} * P_L \tag{2}$$

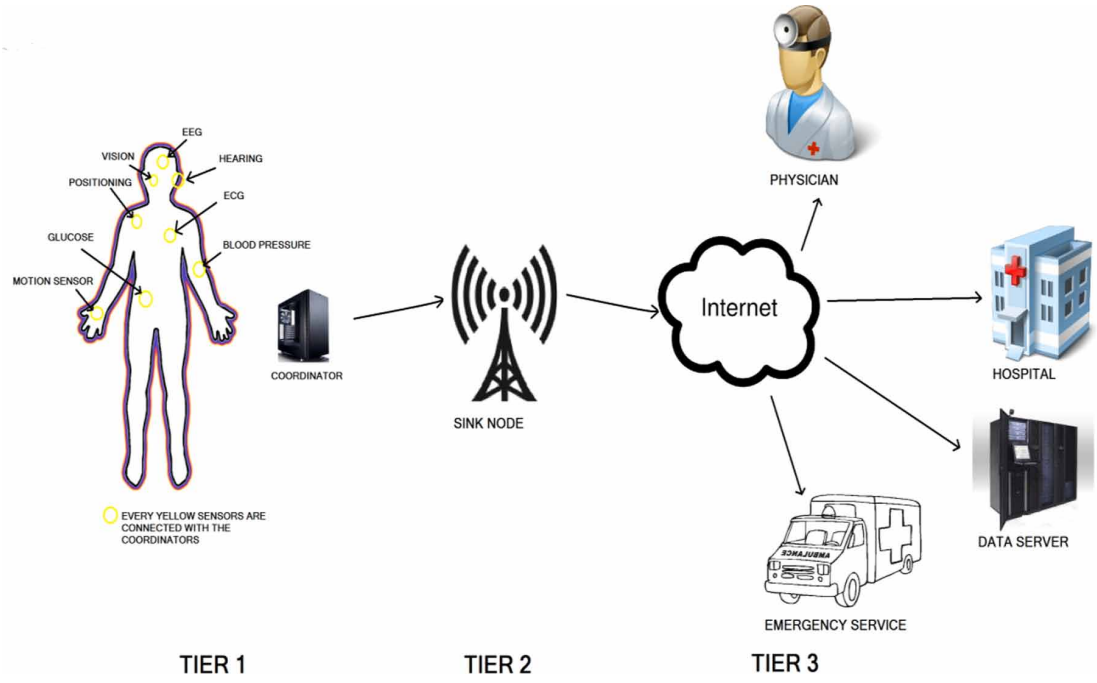
The model takes d n as energy costs due to achannel variation and the path loss in respect with distance d between sender and receiver. Etx represents the transmission energy, Erx the receiver energy, ETX_{elec} and ERX_{elec} the energy the radio dissipates to run the circuitry for the transmitter and receiver, respectively, and Eamp is the energy for the transmitter amplifier. The specific values of these parameters are hardware dependent. In addition, k represents the number of bits sent for the transmission. n is pathloss coefficient related to shadow effect.

The tier 1 of communication in the wireless body area sensor network is based on a number of smart nodes, each accomplished of detecting, sampling, processing, and creating interactive functional signals. Every sensor node obtains initialization guidelines and responds to queries from the personal server. A Coordinator which collects the information from each sensing node and transfers to the SINK node. The tier 2 is the private server that interfaces WBAN sensor nodes, be responsible for the graphical user interface, and communicates with services at the top tier. The private server is classically implemented on a PDA or a cell phone, but instead,it can run on a home private computer. This is predominantly suitable for in-home monitoring of elderly patients. The private server interfaces the WBAN nodes through a network coordinator (NC) that implements ZigBee or Bluetooth connectivity. To interconnect with the medical server, the private server employs mobile telephone networks (2G, GPRS, 3G) or WLANs to reach an Internet access point.

The tier 3, concentrated on a medical server, is enhanced to check hundreds or thousands of discrete users and incorporates a composite network of unified services, medical personnel, and healthcare professionals. Each user wears a number of sensor nodes that are intentionally placed on her/his body. The medical server preserves electronic medical records of enumerated users and delivers several facilities to the users, medical personnel, and informal caretakers. It is the responsibility of the medical server to verify users, receive health monitoring session uploads, design and introduce the particular session data into equivalent medical records, analyze the data patterns, identify serious health variances in order to communicate with emergency caretakers, and forward original directions to the users, such as physician recommended daily workouts. The patient's physician can access the data from his/her workthrough the Internetand observe it safeguard the patient is within probable health metrics (heart rate, blood pressure,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Three tier communications in WBAN



activity), guarantee that the patient is responding to a treatment given by the physician or that a patient has been carrying out the given exercises. A server agent may examine the uploaded data and generate an alert inthe case of a latent medical condition. A larger amount of data collected through these services can also be used for knowledge discovery through data mining. Integration of the collected data into exploration databases and measurable analysis of conditions and patterns could demonstrate the precious information to the researchers trying to link symptoms and diagnosed with chronological changes in health status,physiological data, or supplementaryfactors (e.g., gender, age, weight). It could significantly contribute to monitoring and studying of drug therapy effects for the patients during unavoidable emergency situations.

2.4 Effective Methods

As the total number of sensor nodes in a WBAN is not large, $d_{i,j}$ can be measured manually after the sensor nodes have been deployed as well as p_t, e.g., the transmission path state of each pair of nodes according to the location of each node. Furthermore, the energy consumption parameters are obtained depending on the transceiver types of sensor nodes in the networks. K is identified by the communication protocol and is simply known by WBAN manipulators. E_s can be known by two means depending on two conditions: (A) if the transmit selection is appealed at the network initialization where each sensor node is filled with energy, we can directly achieve the energy storage information from the sensor node form and battery-operated information; (B) if the transmit selection is appealed at a resume from a network recovery, the controller can set a appeal flag in the inspiration frame in order to inform each sensor node to report its residual energy situation in the transmitting time slots. Once all of the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

parameter values are gained, the controller records this information in its memory and has the ability to execute the relay selection scheme. Then, it is illustrated how to implement our relay selection scheme in dual cases. In the principal case where a WBAN is at the initialization stage, the controller will directly execute the relay selection scheme at the beginning of the first frame since it has all values needed to execute the relay selection scheme and load the relay allocation results and the corresponding timeslot allocation in the ideal frame. Then, in the ideal transmission slot, the controller broadcasts the ideal frame to all of the sensor nodes in the network. As each node must listen and receive this ideal frame, all of the sensor nodes in the network will know their transmission strategy and their relay nodes, if needed. In the succeeding case, e.g., a WBAN restarts, where the energy storage condition of sensor nodes is unknown, the controller sets a request flag in the ideal frame at the beginning of the first frame and allocates timeslots for each sensor node in the network to report their energy storage conditions. As a result, in the first frame, sensor nodes send their energy storage information to the controller in their allocated report slots. After the controllers have received all of the report frames, it invokes the proposed relay selection scheme and loads the results and the corresponding timeslot allocation in the next ideal frame, which will be broadcasted in the second frame. Then, in the second frame, each node will know its transmission strategy and their relay nodes. It should be emphasized that the low complexity of the rapid solution in our proposed scheme guarantees that the controller can finish the algorithm in the interval between the point when the controller has received all the energy storage reports to the slot for broadcasting the next ideal frame. Furthermore, it should be noticed that the implementation discussed here does not involve the detailed interaction procedures between the controller and sensor nodes that will be further studied in the future.

3. RELIABILITY

WBANs, which promise significant improvement in the reliability of observing and treating people's health, comprise a number of sensors and actuators that may either be implanted or mounted on the surface of the human body, and which are capable of wireless communication to one or more external nodes that are in close proximity to the human body. Achieving minimal energy consumption, with the required level of reliability is critical for the proper functioning of many wireless sensor and body area networks. Additionally, anyway of the traffic appearances, the methods introduce to understand reliable wireless sensor networks using (occasionally) unreliable components (wireless sensor nodes).

Cooperative communications enable efficient consumption of communication resources, by permitting nodes or terminals in a communication network to cooperate with each other in information transmission. It is a promising technique for yet to come communication systems. The energy efficiency of cooperative systems in wireless body area network (WBAN) to outage the performance of three transmission patterns, specifically direct transmission, single-relay cooperation, and multi-relay cooperation.

At any node *ni*, the reliability estimator module is responsible to measure the average link reliability $\mathrm{LR}_{i,j}$ of link L_{i-j} for neighbor node *nj*. If *N* successful is the number of successful transmissions and *N*total is the number of total transmissions, then the average probability *P*average of successful transmission over a time window δt can be calculated as

$$P_{\text{average}} = N_{\text{successful}} / N_{\text{total}}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. MEDICAL APPLICATIONS OF WIRELESS BODY SENSOR NETWORKS

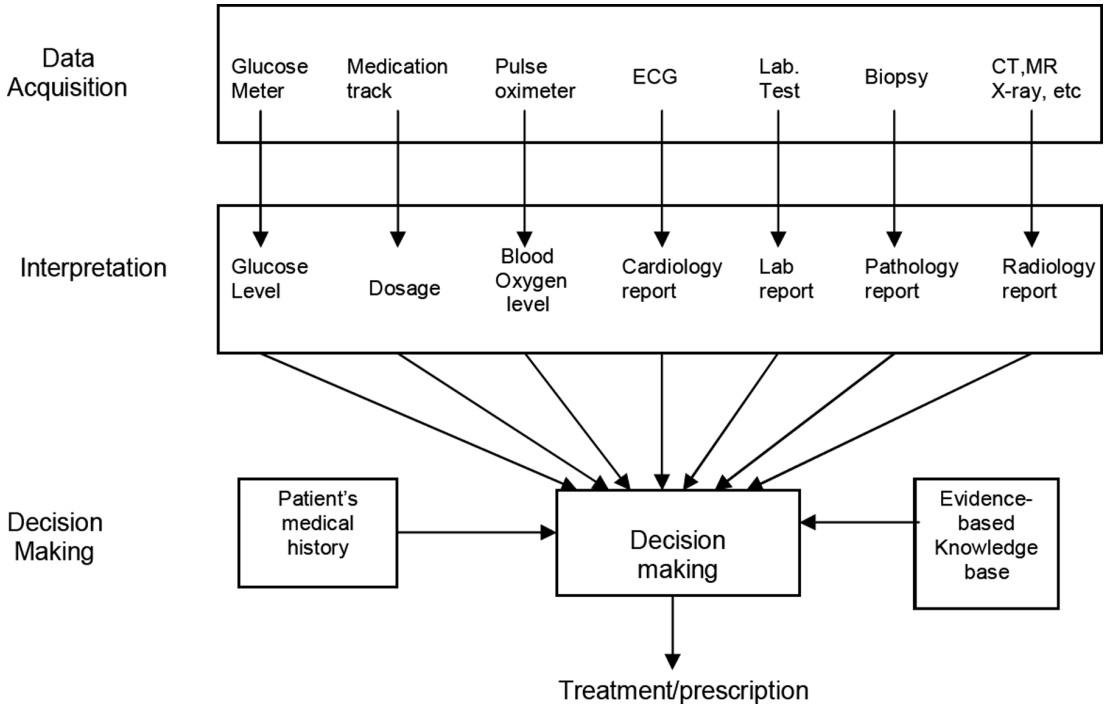
In recent times, the data acquisition begins with the physical property to be measured. Data interpretation and decision making are done based on the data acquisition measured and also treatment is based on the acquisition.

New and improved healing devices in figure 8 are continuously introduced to identify vital signals and present them in an appropriate format for health care supporters. The interpretation can be regarded as a data compression and data conformity process.The general practitioner make a treatment recommendation based on the patient's medical history and current experimental reports by referring the evidence-based database, pharmaco logical handbook, and other resources.

4.1 Energy Consumption and Performance Analysis of WBSN in Healthcare Monitoring

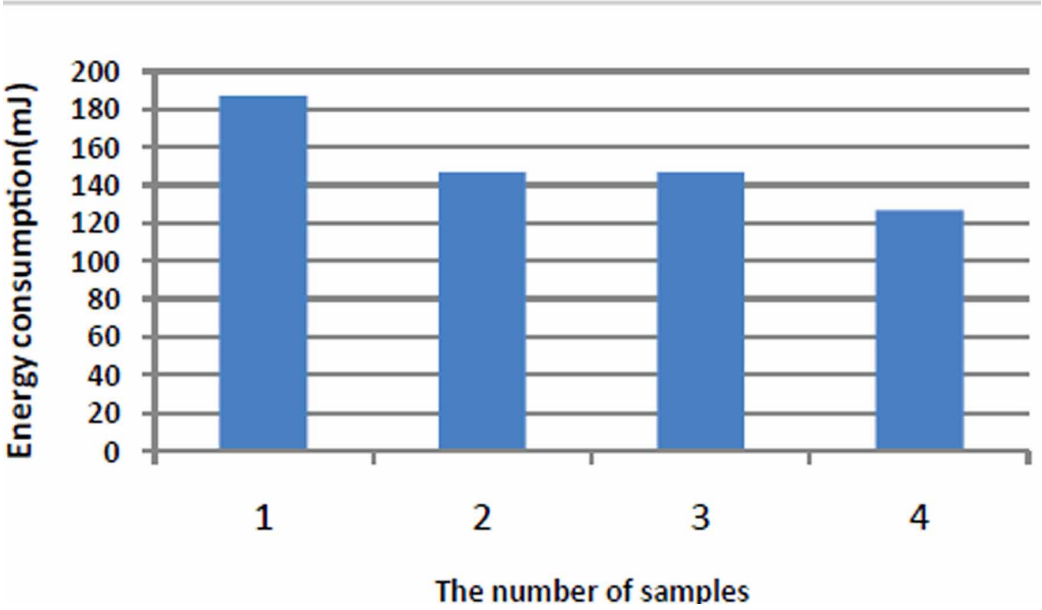
There is a connection between energy consumption and lifetime of the network. A challenge of WBSN is patient eminence of life during the monitoring. So, to provide a quality by reducing energy consumption in a sensor node for the aggregation of sensory data, data packaging and transmitting them to the base station. Therefore, to prolong the lifetime of sensor node when the lifetime becomes short, the patient has to recharge the battery of sensor node in short intervals, and if the patient forget to recharge the battery, it may be dangerous for the patient. Therefore, the burst transmission mechanism is employed to observe the relation between energy consumption and number of samples in a packet.

Figure 8. Applications of WBAN



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Energy consumption for monitoring



4.2 Reliablity of WBSN in Medical Application

The interest in wireless systems for medical applications has been promptly growing. By a number of benefits in wired alternatives, comprises of ease of use, reduced possibility of infection, reduced the possibility of failure, reduce patient uneasiness, improve mobility and low cost of precaution conveyance, wireless applications convey exciting potentials for new applications in the medical market. Portable devices such as heart rate monitors, pulse oximeters, spirometers and blood pressure monitors are essential instruments in intensive care. Integrated wireless technology, these wireless devices could communicate with a gateway that connects to the medical center's network and transmits data to health data stores for monitoring, control, or evaluating in real time or offline after storage.The WBSN provides a reliable updates about the patients to the doctors by using any wirless medium using network routing protocols. The protocols avoides delay and packet loss during transmission

5. CONCLUSION

The essentiality of Lifetime Enhancement and Reliability in Wireless Body Area Network is discussed in this paper.Which describes types of WBAN and the usage of it based on our requirement and patientsproblem. Then the transmission path state of each pair of nodes is located according to the location of each node. Moreover, the energy consumption parameters are acquired depending on the transceiver types of sensor nodes in the networks. Reliability is maintained by the cooperative communication between the nodes. Thus a prolonged network lifetime achieves a reliable network with minimum delay and maximum throughput while considering power consumption by reducing unnecessary communication.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6. REFERENCES

Samaneh, M., Mehran, A., Justin, L., David, S., & Abbas, J. (2014). Wireless Body Area Networks: A Survey. IEEE Communications surveysand tutorials, 16(3), 1658-1686.

Riccardo, C., Flavia, M., Ramona, R., Chiara, B., & Roberto, V. (2014). A Survey on Wireless Body Area Networks: Technologies and design challenges. IEEE Communications surveys and tutorials, 16(3), 1635-1657.

Emmanuel, D., Kola, S., & Mohana, J. (2014). A Survey on wireless Body Area Networks. International journal of scientific and research publications, 4(3), 1-7.

Conroy, L., Ó Conaire, C., Coyle, S., Healy, G., Kelly, P., Connaghan, D., ... & Nixon, P. (2009). TennisSense: a multi-sensory approach to performance analysis in tennis. In *Proceedings of the 27th International Society of Biomechanics in Sports Conference 2009*, Limerick, Ireland, August 17-21.

Smail, O., Kerrar, A., Zetili, Y., & Cousin, B. (2016, September). ESR: Energy aware and Stable Routing protocol for WBAN networks. In *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)* (pp. 452-457). IEEE.

Hadda, E. B., Lamia, C., & Lotfi, K. (2012). A Survey of routing protocols in Wireless Body Area Networks for healthcare applications. International journal of Ehealth and medical communications, 3(2), 1-18.

Kavitha, D., Balapriya, M., & Sundrarajan, T. V. P. (2016). A Survey of routing protocols in Wireless Body Area Networks. *South Asian Journal of Engineering and Technology*, *2*(21), 44–51.

Ahmad, J., & Zafar, F. (2012). Review of body area network technology & wireless medical monitoring. *International Journal of Information and Communication Technology*, *2*(2).

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). A Survey on Sensor Networks. *IEEE Communications Magazine*, *40*(Aug), 102–114. doi:10.1109/MCOM.2002.1024422

Patel, M., & Wang, J. (2010). Applications, challenges, and prospective in emerging body area networking technologies. IEEE Wireless communications, 17(1), 80-88.

Khan, J. Y., & Yuce, M. R. (2010). Wireless body area network (WBAN) for medical applications. In New developments in biomedical engineering. InTechOpen.

Crosby, G. V., Ghosh, T., Murimi, R., & Chin, C. A. (2012). Wireless body area networks for healthcare: A survey. *International Journal of Ad Hoc. Sensor & Ubiquitous Computing*, *3*(3), 1.

Pervez Khan, M., Hussain, A., & Kwak, K. S. (2009). Medical applications of wireless body area networks. *International Journal of Digital Content Technology and its Applications, 3*(3), 185-93.

Carrano, R. C., Passos, D., Magalhaes, L. C., & Albuquerque, C. V. (2014). Survey and taxonomy of duty cycling mechanisms in wireless sensor networks. *IEEE Communications Surveys and Tutorials*, *16*(1), 181–194.

Reddy, G. P., Reddy, P. B., & Reddy, V. K. (2013). Body area networks. *Journal of Telematics and Informatics*, *1*(1). doi:10.12928/jti.v1i1.36-42

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

González-Valenzuela, S., Liang, X., Cao, H., Chen, M., & Leung, V. C. (2013). *Body area networks. In Autonomous Sensor Networks* (pp. 17–37). Springer.

Ben Elhadj, H., Chaari, L., & Kamoun, L. (2012). A survey of routing protocols in wireless body area networks for healthcare applications. *Int. J. of E-Health and Medical Commun.*, *3*(2), 1–18. doi:10.4018/jehmc.2012040101

FCC. (2015). Medical Body Area Networks. Retrieved from http://www.fcc.gov/document/medical-body-area-networks

Seyedi, M., Kibret, B., Lai, D. T., & Faulkner, M. (2013). A survey on intrabody communications for body area network applications. *IEEE Transactions on Biomedical Engineering*, *60*(8).

Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D., & Jamalipour, A. (2014). Wireless body area networks: A survey. *IEEE Communications Surveys and Tutorials*, *16*(3), 1658–1686. doi:10.1109/SURV.2013.121313.00064

Khan, Z. A., Sivakumar, S., Phillips, W., Robertson, B., & Javaid, N. (2015). *QPrd: QoS-aware peering routing protocol for delay-sensitive data in hospital body area network*. Mobile Information Systems.

Misra, S., & Sarkar, S. (2015). Priority-based time-slot allocation in wireless body area networks during medical emergency situations: An evolutionary game-theoretic perspective. *IEEE Journal of Biomedical and Health Informatics*, *19*(2), 541–548. doi:10.1109/JBHI.2014.2313374 PMID:24686307

Tang, Q., Tummala, N., Gupta, S., & Schwiebert, L. (2005). Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue. *IEEE Transactions on Biomedical Engineering*, *52*(7), 1285–1294. doi:10.1109/TBME.2005.847527 PMID:16041992

Binkley, P. (2003). Predicting the potential of wearable technology. *IEEE Transactions on Biomedical Engineering*, *22*, 23–24.

Otto, C., Milenkovic, A., Sanders, C., & Jovanov, E. (2006). System architecture of a wireless body area sensor network for ubiquitous health monitoring. Journal of mobile multimedia, 1(4), 307-326.

Khan, J. Y., & Yuce, M. R. (2010). Wireless body area network (WBAN) for medical applications. In C. Domenico (Ed.), *New Developments in Biomedical Engineering* (pp. 591–627). INTECH publishing.

Filipe, L., Fdez-Riverola, F., Costa, N., & Pereira, A. (2015). Wireless body area networks for healthcare applications: Protocol stack review. *International Journal of Distributed Sensor Networks*, *11*(10), 1–23. doi:10.1155/2015/213705

Wu, T., Wu, F., Redouté, J. M., & Yuce, M. R. (2017). An autonomous wireless body area network implementation towards IoT connected healthcare applications. *IEEE Access : Practical Innovations, Open Solutions*, *5*, 11413–11422. doi:10.1109/ACCESS.2017.2716344

Negra, R., Jemili, I., & Belghith, A. (2016). Wireless body area networks: Applications and technologies. *Procedia Computer Science*, *83*, 1274–1281. doi:10.1016/j.procs.2016.04.266

This research was previously published in the International Journal of User-Driven Healthcare (IJUDH), 8(2); edited by Rakesh Biswas and Ashok Kumar Biswas; pages 46-59; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 4
Utilization and Applications

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 51
Pluralistic Tendencies in Healing Abantu Illness in the Contemporary South Africa

Kholekile Hazel Ngqila
Walter Sisulu University, South Africa

ABSTRACT

Ukuhanjwa illness was used as an example to understanding abantu illnesses. With attributional theory ukuhanjwa illness is attributed to spiritual and social causes rather than biomedical causes, whereby causal link is socially constructed between ukuhanjwa illness and entry into the body by familiars. Issues explored included conceptualisation of ukuhanjwa illness. The focus of the chapter is on the reasons for continued pluralistic tendencies in healing regardless of the expectation by the West that people should be focusing on the use of the fast evolving biomedical healing methods. The ethnographic study took place among the Southern Nguni people of OR Tambo District Municipality (ORTDM) in the Eastern Cape, South Africa. Data was collected using qualitative and ethnographic research methods amongst a sample group of 50 participants. The sample was composed of traditional healers, mothers of children who have experienced ukuhanjwa illness, elderly people (male and female), biomedical practitioners and nurses.

INTRODUCTION

The chapter begins by presenting the background for the discussion by outlining the description and identifiable symptoms of *ukuhanjwa* illness; aetiology of *ukuhanjwa* illness; and dealing with illness and disease in order to have better understanding of the illness in question. The chapter has also outlined the focus of the discussion as the rationale for the Southern Nguni people to continue to engage in pluralistic tendencies in healing *ukuhanjwa* as one of the *abantu* illnesses regardless of the expectation by the West that people should be focusing on the use of the fast evolving biomedical healing methods. Ngubane (1977) in Urbasch (2003, p. 11) defined *abantu* illnesses as spiritually caused illnesses which can only be healed using traditional healing methods. Pluralistic approach has been found to be popular

DOI: 10.4018/978-1-7998-1204-3.ch051

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

even in other developing countries such as India (Whiting, 1977, p. 249). Dayes (2014) has also noted that people resort to pluralistic tendencies in healing medically unexplained symptoms. The chapter reveals how the Southern Nguni people attach different interpretations and meanings to their illnesses and health experiences as one rationale for people to engage in pluralistic tendencies in healing as their response. The chapter specifically looks deeply into the phenomenon of invisibility of *ukuhanjwa* to the biomedical health system within the context of power relations between the two medical systems – the biomedical and the traditional as another rationale for the use of pluralistic approach in healing *ukuhanjwa* illness. Invisibility has both literal and contextual meanings in relation to *ukuhanjwa*. The invisibility of *ukuhanjwa* to biomedical diagnosis has also been addressed as part of a discourse which needs to be viewed either within its historical, political or cultural setting (Lupton, 2012, p. 2). The chapter has mainly focused on the cultural setting. This discourse also includes the concept of ritual purification, for without such discourse the term is apt to be seen as directly contradicting the literal meaning of the word - pure (Good, 2001, pp. 92-103).

Pluralism is often the result of beliefs that certain illnesses are invisible when using biomedical means of diagnoses. The Southern Nguni people also view pluralistic healing method as the holistic approach in healing *ukuhanjwa* illness – hence this has been discussed under rationale for pluralistic tendencies in healing. The chapter also addresses the perceived source of the healing power as a powerful motivation for continued use of traditional healing methods for *ukuhanjwa* and how it can be linked to pluralistic tendencies in healing.

Lastly, the chapter presents solutions and recommendations highlighting the implications of the Southern Nguni beliefs and cultural practices regarding health, illness and medicine and how those implications relate to the biomedical health sector domination as well as attempts to introduce social change.

BACKGROUND

Description and Identifiable Symptoms of Ukuhanjwa Illness

The Southern Nguni people of OR Tambo District Municipality (ORTDM) in the Eastern Cape of South Africa describe *ukuhanjwa* as an attack on a person, young or old, by familiars (*izilwanyana zokuthakatha*), literally meaning demonic animals used by witches, which penetrate the body through any bodily opening, resulting in certain identifiable and literally visible conditions. These identifiable and literally visible conditions become contradictory to the invisibility which will be explained in the discussion to follow as a contributing factor to the continued pluralistic tendencies in healing what is believed to be *abantu* illnesses.

To name the illness is to diagnose the illness. Once the name is known, the symptoms are understood, the cause is known and the treatment is straightforward. To the Southern Nguni, the illness is known as *ukuhanjwa* from *uku* meaning – to and *hanjwa* meaning - penetrate. Thus, literally - to penetrate and idiomatically - spiritual penetration. This name arose, as do all names for illnesses, from shared experiences and the common worldview of a particular group. The Southern Nguni people make a distinction between two forms of *ukuhanjwa*, namely, the mild and the severe forms of the illness. The mild form manifests itself as red patches on the lips, around the outer lining of the anus and the genitals, as well as a wide open anus. The victim could be a child or an adult since *ukuhanjwa* has no age limit, but is usually a child or infant. The victim is always sleepy and, in the case of babies, their cry is always indicative

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of weariness rather than of pain or discomfort. Indeed, anecdotal evidence suggests that the mild form can present with no pain at all. For Southern Nguni people, only a loud scream is indicative of illness with pain. This form is fairly easy to control and is usually dealt with by the mother or caregiver using traditional medicines.

The severe form of the illness attacks the upper parts of the body as well as the lower parts, with the sinking of the fontanel, the neck losing muscle strength, regurgitation of feeds (*uxakaxa*) and thick saliva from the mouth. Vomiting and diarrhoea are common at this stage. The anus tends to be particularly wide open, to the point where one is able to see up into the intestines. The genitals may also be open. There could be a ring of red patches around the anus and genitals and pus may emit from both the nose and anus of the victim.

The Southern Nguni people of the ORTDM express their fear of severe *ukuhanjwa*, stating that, since it starts in the upper parts of the body with symptoms that could indicate a variety of illnesses, it is very difficult to identify. The most common response of the Southern Nguni when they suspect *ukuhanjwa* is to check the anus of the victim; that is, whether or not it is wide open. They say that, once the victim reaches the stage of losing muscle strength in the neck, there is no way of recovering, and the victim will die.

Aetiology of Ukuhanjwa Illness

The attributional theory as explained by Vaughn, Jacquez and Baker (2009, pp. 64-67) is relevant for contextualising the Southern Nguni conceptualisation of the spiritual causation of *ukuhanjwa*. This illness is attributed to the behaviour of familiars acting on their own, without being instructed by witches. Instead, penetration of the familiars into the bodily openings of the victims is believed to be as a result of their experience of hunger, having nothing to do at that particular moment of their penetration or very hot weather conditions according to the explanation given by the respondents. Another person could argue that although the familiars are believed to cause *ukuhanjwa* illness without the witch's volition, the mere fact that the familiars which have harmed the victim are possessed by a particular witch – that which should be held responsible for all the evil acts of his/her familiars.

Contrary to that thought, the Southern Nguni people consider *ukuhanjwa* illness to be different from other *abantu* illness conditions. Victims regarding *ukuhanjwa* illness are also understood to be blameless with less control over the condition (Foster, 2009, p. 116). Instead people become more concerned about urgent healing and recovery of the victim without worrying about who has caused the illness. In the case of other *abantu* illness conditions witches are blamed for instructing their familiars to cause harm to the victims as a result of enmity or jealousy. In the western, scientifically-defined rationalistic framework, witchcraft might not be recognized as real, but within African communities, it is a reality for most people (Petrus & Bogopa, 2007, p. 2). This points to Barker's (2003, pp. 47-60) statements regarding knowledge and truth. Barker (2003, pp. 47-60) states that knowledge is a matter of perspective, not pure or neutral - attained from a particular point of view. Hence the claim that truth can be viewed as a social commendation rather than an accurate picture of an independent reality, stemming from cultural pluralism as a way of generating shared meanings by different cultural groups (Barker, 2003, pp. 51-60). This conflicts with the biomedical view of itself as representing absolute truth based on empirical evidence – empirical evidence defined from the western perspective. According to Petrus and Bogopa (2007, p. 2), both urban and rural African communities believe that witchcraft exists as a viable and plausible cause for misfortune, illness and death, regardless of whether these things can be scientifically explained or not.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This understanding is applicable to familiars commonly associated with men, such as a certain type of snake called *ichanti*, as well as familiars associated with women, such as a certain type of mouse called *impaka*, or a snake called *irhamba* (puff adder). According to the informants, *ichanti* is known to be invisible to people in its snake form, but takes the form of anything else – such as a horse, or water. Hence the informants could not describe it, other than stating that it is a snake which helps the owner to accumulate wealth. The description of *impaka* indicates something between the *Graphiurus murinus* (woodland dormouse) which has a bushy tail, known to be endemic to South Africa and the *Elephantulus edwardii* (Cape elephant Sengi) which has a very long nose. *Impaka* has both a bushy tail and a long nose, and therefore falls somewhere between the two descriptions. These are the familiars which are believed to be responsible for causing *ukuhanjwa* illness.

Dealing with Illness and Disease

How a person responds to illness and disease depends entirely on his or her worldview. A worldview may be shaped by many factors, some of them contradictory, and may result in an approach which we would call pluralistic. This may be seen as an attempt to reconcile apparently opposing views. People tend to use either traditional, or pluralistic or biomedical healing approaches. Pluralistic healing approach is when the ill-person prefers to use traditional and biomedical healing approaches simultaneously or starts by using traditional healing approach and later use biomedical healing approach as a last resort or starts by using biomedical healing approach and later use traditional healing approach as a last resort. *Ukuhanjwa* is so unique a condition that it tends to be identified as a case for spiritual remedies (traditional healing approach) first and foremost.

One of the most striking aspects of *ukuhanjwa* illness is that, despite the fear which surrounds it and the seriousness of the sickness, the Southern Nguni have developed no particular strategy to pre-empt it. All responses are at the point of attack and immediately afterwards. No long term measures are adopted to prevent it from re-occurring. This seems to be chiefly because the illness is considered fairly rare, and strategies that would probably involve the daily application of an herbal concoction are deemed unnecessary. Apparently, it is not that there is no way of preventing occurrences of *ukuhanjwa*; indeed the same medicines used to cure it could be used to keep it away for the long term. The reality is that people forget, and do not feel a sense of urgency about preventative measures, particularly as these attacks are not common.

The strong smell of the medicines is considered unbearable to the familiars, and is used to repel them, but for a short period of time, just long enough for the person to heal. As soon as the victim is healed, use of the medicine for chasing away familiars is discontinued. The possibility could be that the periods between attacks are just too long, if they recur at all, and the time period is just too long to maintain consistency about the preventative measures. Since it is believed that familiars often loiter around waiting to be summoned and used by their owners, they could be anywhere at any time, and it would imply a very vigilant and daily use of the medicines in order to be sure that one has kept them away.

Regarding cures for *ukuhanjwa*, there are specific traditional remedies, and these seem to work perfectly well – at least in the case of mild *ukuhanjwa*. A teaspoon of a concoction made from the leaves and/or roots of *isivumbampunzi* (*Tulbaghia violacea*), or the leaves *iboza* (*Plectranthus barbatus*) is given as a drink to the victim, and droplets of the same concoction are administered into all bodily openings. For severe *ukuhanjwa*, the Southern *Nguni* people take their guidance from the diviner. In this case it is always a decoction made from the boiled familiar which has attacked the victim. Familiars associated with female witches are the mouse (*impaka*), and the puff adder (*irhamba).*

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In cases where the attack is by *ichanti*, the familiar used by male witches, a special ritual known only to the male owner of this snake needs to be performed. Usually the owner gets to hear of the illness and comes forward voluntarily to offer to heal the child. But the healing has to be performed privately in the presence of the male owner of the familiar and the *ukuhanjwa* victim and in the absence of the caregiver. The involvement of the male owner of the familiar is only at the severe stage of *ukuhanjwa* illness. In consideration of the sample size for the research, chances for survival at the level of severe *ukuhanjwa* illness proved to be less than 50% as out of three participants only one had a child who was successfully healed.

Southern Nguni people are very specific about the ritual space and time of day for harvesting the indigenous medicinal plants. These indigenous medicinal plants need to be ritually pure and have the required maximum healing power, and for this, the place and time for collecting the herbs must be revealed to the diviner by his or her ancestors, through dreams or through consultation of the ancestors by the ordinary person via the diviner. Additionally, collection of indigenous medicinal plants within the premises of the homestead for healing *ukuhanjwa* illness has never been encouraged as the Southern Nguni belief is that their ritual purity and healing power can be weakened by the familiars by walking over them before and after the attack of the victim. The same applies for any person who happens to be treating the ill person - significance of ritual purity of the person performing healing as well as the plants to be used for healing, ritual timing and ritual space have to be highly honoured and observed. The whole healing process of *ukuhanjwa* ends up being ritualised, and no one has ever wanted to deviate from the stipulations that already exist for curing this illness.

RATIONALE FOR PLURALISTIC TENDENCIES IN HEALING UKUHANJWA

Interpretations and Meanings of Illness Experiences

The interpretation of a particular incident or illness depends on the social construction of illnesses in general. An example can be made of an elderly woman who participated in interviews on *ukuhanjwa* when she was 72 years of age. She was a well-educated person with a degree in education and a retired school teacher. She told of how she had developed a small pimple on her leg which she did not take seriously until it developed into a wide circle of black pigmentation. She had not experienced any pain and to her this was significant, and an indication of witchcraft; it being the intention of witches that she would disregard the wound until it caused complete dysfunction of the leg. Her expectation was that a natural illness would be accompanied by pain; the lack of pain was mysterious, indicating the witches' strategy to mislead her so that she would not take action. When advised to visit the biomedical health centre, she refused giving her own interpretation of the cause as a spell she walked over which was set for her specifically by the witches at her younger age so that she would be harmed and fall ill.

Based on her interpretation of the cause, her conclusion was that the biomedical health centre would surely attach an incorrect name to the illness and attempt to heal her, but would not succeed. The true nature of the illness would be medically invisible to biomedical diagnosis.

Her interpretation of the illness led her to decide to apply indigenous medicine to her wound without hesitation. She used the leaves of the sweet potato. She crushed the leaves to extract the juice which she applied directly to the wound on a daily basis for about a week, causing the wound to burst open and release a black fluid. Release of the fluid brought healing. Her interpretation of the experience was that

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the juice of the sweet potato assisted in releasing the spell in the form of the black pigmented fluid. That the wound got healed after the application of the medicine confirmed, in her mind, her initial diagnosis and interpretation.

Invisibility of Ukuhanjwa to the Biomedical Health System and the Southern Nguni Contextual Meaning of Invisibility

According to Liebeck and Pollard (1995, p. 897), visibility refers to being able to be seen or noticed. Applying the term 'invisible', to the symptoms of *ukuhanjwa* could be confusing to the outsider as the symptoms are clearly visible. But the term refers to the whole condition, not only to its symptoms – its causes are included in the term. The biomedical interpretation of the condition varies but includes the suspicion of sexual assault, nappy rash and diarrhoea. The loss of muscle strength in the neck and the sunken fontanel are seen as results of severe dehydration.

When a caregiver takes a child suffering from *ukuhanjwa* to a biomedical centre, she is immediately faced with problems as the tendency is for a biomedical practitioner to separate the cause from the symptoms and see little correlation between the symptoms. The caregiver would be left feeling dissatisfied with the explanation rendered by the biomedical health sector on the illness causation and method of healing that they would receive. Sometimes the victim or the caregiver may just mention the name of the illness *ukuhanjwa* to the biomedical nurse or practitioner without giving a full description of the symptoms. The biomedical practitioner would find the name meaningless as it is not a recognised condition.

In indigenous healing, purity, too, is something that is present but not seen. According to the Southern Nguni people, all indigenous healing methods which deal with spirituality and spiritually-caused illnesses are believed to be ritually pure only if their use lead to successful ritually purification and healing of the ill or polluted person or successfully prevent pollution before it occurs. This is one of the reasons why people who believe in *ukuhanjwa* cannot easily give up their indigenous healing methods – ritual purity is of fundamental importance to them. The very strong smell and taste of indigenous and indigenised medicines used for healing *ukuhanjwa* are believed to be evidence of their ritual purity. The taste and smell contain the power to ward off pollution and to prevent familiars from further polluting the same victim until the process of healing has been successfully achieved. The ill person is not given the medicine for purging but it is supposed to remain inside for cleansing the dirt/impurities from the inside to the outside in order to cool the ill person. According to the Southern *Nguni* people, the contextual meaning of *ukungcola* (dirt/impurities) and *ubushushu* (heat) is pollution. They are the tracks (*umkhondo*) or traces which are left by the penetration of the familiars, and they require ritual purification for cleansing in order to achieve successful healing.

The bitter and biting taste of the indigenous medicines used for the illness in question is believed to have the ability to magically squeeze out pollution from the afflicted parts of the body by means of sympathetic magic. The result is that bodily openings that were harmed and deformed return to their original size. Leclerc-Madlala (2002, p. 93) explains that sympathetic magic is a situation whereby medical conditions are believed to be treatable by substances that are symbolically associated with the condition. It is believed that if the medicine is able to cause the facial muscles to squeeze, then the same response is bound to happen with the bodily openings which have been harmed by the familiars to magically heal and restore their original shape and size.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Holistic Approach in Healing *Abantu* Illnesses

A research report by Ross (2010, p. 46) which stated that for every ten black people, eight use traditional medicine alone or in combination with biomedicine was confirmed in this study as true for the Southern Nguni people of ORTDM. A strong factor influencing the pluralistic nature of help seeking is that people wish to avoid the accusation of not having really tried to find a cure. Since illness is viewed holistically, as spiritual as well as physical, all avenues must be tried.

Another reason mentioned by Eastman (2011, p. 184) is that traditional healers are more easily and widely available as an integral part of the population's social and cultural system. This might not be true for the Southern Nguni people of ORTDM as their behaviour is influenced by their belief system far more than the availability or non-availability of both medical systems. Many urban people, it has been pointed out, will travel for miles to find the right diviner. And many of these same people live in close proximity to biomedical centres.

These pluralistic tendencies are influenced by the belief that healing is not just about the physical body only, but also about spiritual healing – hence, for the Southern Nguni people of ORTDM, the belief system is the chief factor in deciding where to seek help. As Kleinman (2010, p. 86) says, different medical systems employ different explanatory models in an attempt to make sense of disease and give meaning to the individual and social experience of illness. Thus many people will explore all options at their disposal so as to minimize the blame for ill-health or death, since according to their conceptualisation illness and health are defined holistically.

Another reason for the pluralistic tendency in healing is that what is perceived as illness in one culture may not be perceived the same way in another culture. Witchcraft is not recognized in the biomedical health system while for traditional people it is a reality – hence traditional people are open to all options for healing at their disposal.

Eastman (2011, p. 184) confirms that traditional healing together with biomedical healing is widely practiced across social boundaries, as it is not confined to a particular generation or even socio-economic standing. One notices a particular pattern amongst urban people who are assumed to be acculturated. They begin with the biomedical means of healing and later resort to traditional healing methods. Some would immediately engage with both. Their reason for doing this would be their belief that even if the illness starts by being physical, evil spirits or familiars are opportunistic and will attack the victim when he is weakest, so that the ill person confuses spiritual with physical illness. This has also been realised by Farrand (1984, p. 2) and Shoko (2011, p. 290) although Shoko goes on to mention that this pluralistic nature of healing is not only prominent with the urban black people but is also found in the rural areas.

Thinwa (2004) notices the same pattern with urban Tuberculosis (TB) and HIV/AIDS patients, where about 70% adopt a pluralist approach in an attempt to overcome their fear and recover as quickly as possible. Bhana (1986, p. 221) had a similar encounter with Indians who had the same practice of seeking help from both biomedical centres and traditional healers, especially for mental illnesses. The Chinese have established that the same illness can be healed using different methods of healing (Hu, Du, Shen & Xu, 2012, p. 2). Thus the pluralistic tendencies around healing are common and widespread.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Power Relations and Discourses in Illness, Health, and Medicine

Illness, health and medicine are politicised fields in this era of a huge concentration of power in the hands of few. Globalisation has ensured the dominance of a particular worldview and healing system that is inextricably bound to political and economic factors. Hence the interest of medical anthropologists in how health and medical language relate to power (Lupton, 2012, p. 12).

The nature of diseases and their cures, and the process of identifying and labelling diseases, has been dominated by the biomedical health system to the point where it is unable to recognise any other interpretations but its own. It is unable to recognise, name and cure *ukuhanjwa*. Faced with this reality, it strives to treat symptoms, but not the disease itself. The biomedical response to *ukuhanjwa* symptoms is to treat each one individually. They would see the wide open anus as a result of diarrhoea or sexual assault, and advise accordingly. Usually they dispense paracetamol or acetaminophen (the generic name) under the trade name Panado, to bring down the temperature where the *ukuhanjwa* is mild. The traditional approach is to get to the root of the problem and deal with all symptoms at once. It is a truly holistic approach.

According to the Southern Nguni people, biomedical practitioners do not share the name '*ukuhanjwa*' for the condition because they do not recognise the cause as a valid one. To them it is a collection of symptoms without a name – a non-illness. As a non-illness, it cannot be treated effectively. The Southern Nguni way of interpreting the non-comprehension of biomedical practitioners of this condition is that its true nature is invisible to them. They observe that biomedical practitioners cannot provide a medicine or healing method which will simultaneously heal all the symptoms and chase away the evil spirits. According to the Southern Nguni people, the western healing methods heal the physical body whereas traditional healing methods target even the evil spirits, extracting the illness from the deepest parts of the body. Traditional healing methods work on the causes of the illness as indicated that it targets the evil spirits for holistic healing. The enema has to be a small amount, not for purging the victim, but to allow it to cleanse his/her internal organs so that when releasing the bowels it comes out together with the evil spirit and the heat, which resulted from the contact with the familiars, for the victim to cool.

What stands out in this method of healing is that there is no attempt to focus on this or that particular symptom; it is always seen as a collective of symptoms, healed simultaneously. Seemingly, it was originally important for the Southern Nguni people to get the biomedical health sector to understand their illness, and it troubled them that biomedical practitioners continued to dispute its existence. Hence they resorted to the belief that the illness is invisible to the biomedical world. Comparisons between the two approaches would be quite unfair, as they arise from such totally different worldviews. And yet with this particular disease, it does seem that traditional medicine has the upper hand. It has the record of curing from the contextual meaning of the Southern Nguni people, whereas the biomedical approach does not.

Linguistics in medicine, of course reflect the culture from which the medical system arises. The way people speak or the visual metaphors they use often cause confusion, misconceptions and misinterpretations. The pattern of words, figure of speech, concepts, values and symbols (Douglas, 2010, pp. 9-10; Lupton, 2012, p. 2) all contribute to different meanings attached to phenomena, as the case with the phenomenon of cleanliness or purity. The Southern Nguni concept of ritual purity has sometimes erroneously been equated with hygienic cleanliness (Ngqila, 2002), resulting in misunderstandings. To the Southern Nguni, purity and cleanliness in medical matters refer to spiritual cleanness, and the use of anything which brings about this state of spiritual cleanness. This includes appeasing ancestors, warding off pollution and healing the ill. It does not mean sterile, free of germs, clean in the physical sense.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Thus we see that different cultural groups develop ways of making sense of illness, health and medicine from their own cultural perspectives, based on historical, political, economic and cultural settings which cannot be compared, as already stated. As a way to rationalise the great disparity that exists between their own traditional methods and biomedical methods, traditional Southern Nguni accept that some healing methods specialise in physical healing while others specialise in spiritual healing. This allows a place for both.

The manifestation of power relations in the field of health and illness can be seen in the way a stronger culture has come to dominate the discourse on health and healing. According to one of the participants (a community member and a nurse), one of the strategies used by the South African government as a way of overpowering those who subscribe to traditional healing methods has been to introduce laws forbidding people from withholding treatment from biomedical health centres to children. This participant acknowledged that adults are still allowed a choice but it is always insisted that they should allow four hours after the use of biomedicine before the application of indigenous medicine. The understanding is that four hours is the lifespan for the functioning of the modern treatment. The biomedical sector does not accept use of traditional medicine at all, such that it is always made to be clear that if a person decides to use traditional medicine, she or he would be using it at her or his own risk and should not blame the biomedical centre. But when it comes to children, the participant insisted that it is a requirement of law that every means possible be made to cure the child biomedically. But it turned out that the protection of children's rights to exclusive medical care (biomedical care) is merely theoretical as, according to the same participant, a particular parent collected her child from the local hospital in preference for traditional medicine because it was after a long while that the hospital had been trying to heal the child without success. The mother of the child arrived at the hospital stating that her child has been attacked by the familiar called *ichanti* (a male snake believed to be for accumulating wealth) and that someone has offered to heal the child and she asked for permission to leave with the child. She explained that the person who offered to heal her child was the owner of the *ichanti* snake. The hospital released the child because the child could not be healed biomedically for a very long time.

The incident confirms what was stated by Mofokeng (2003, pp. 15-16), that some medical practitioners are themselves confused what to advise when they have been exposed to a variety of cultures and experience cultural conflict. In this regard, the nurses had a responsibility to abide by the rules of biomedicine, but because they have been exposed to traditional practices too, they found themselves bending the biomedical rules and releasing the children to their parents. As Levine (2012, p. 86) says, even though biomedicine is accepted as authoritative, its knowledge is not always uniformly shared and practised by all its practitioners. This leads to different diagnostic and therapeutic possibilities, as in the case above (Levine, 2012, p. 86).

This leads to the consideration of ethnocentrism and relativism. Although lip service is paid to the ideas of relativism, with biomedical practitioners encouraged to understand the so-called inferior and weaker cultures, in reality the ethnocentric approach of biomedicine usually swallows up all other considerations. It has already been noted by one of the participants (a community member and a nurse) that the biomedical health centres are clear about what they want to achieve in terms of bringing about social change in the area of illness, health and healing. They would like the wholehearted adoption of biomedical healing methods without any pluralistic tendencies. There are however exceptions in private practice. During interviews in ORTDM, the researcher was told of two biomedical practitioners who attracted a large following by adapting their own healing methods to the needs of people presenting with *abantu* illnesses. These two biomedical practitioners, one a black African man and one a white man from abroad, started working together in Tsolo and later in Mthatha.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

People would travel from all over the ORTDM to these two doctors, because they showed unusual sensitivity to patient's worldviews and understandings of their own conditions. They listened carefully to the language of the local people over a number of years and were able to pick up the terms they used for illnesses. They then indigenised certain biomedicines such as *udupha* (doepa) or *ivimbela* to suit the needs of the traditional people. One of the interviewees from Mhlontlo Local municipality MLM (Tsolo area) under the ORTDM mentioned that she took her children to these practitioners whilst they were still working jointly in Tsolo for all *abantu* illnesses such as *ukuhanjwa,* and their treatments worked. When the researcher visited one of these biomedical practitioners who is currently practising alone in Mthatha, he told of how traditional people who came to him for treatment would adopt the practice they used when consulting diviners: they would sit in front of him and not say a word. His assistant had to explain this to him that when traditional people visit diviners they do not talk, instead it is the diviner who talks, stating the nature of the illness of the patient.

This inspired the biomedical practitioner to purchase a diagnostic machine which takes the place of the diviner as a scanner. The patient sat on the chair and the chair would be able to detect and indicate possible illnesses such as the level of blood pressure, diabetes and others without the patient or practitioner having to say anything. The patients ended up naming the machine as *igqirha* (the diviner). But unfortunately, the biomedical practitioner avoided mentioning the name of the machine and seemingly he is still the only one who has the machine in Mthatha. Even when the researcher tried to inquire about the name of the machine from other biomedical practitioners, they could not tell as they indicated that they have never heard about such machine. The researcher saw the machine which resembled the chair where the patient could sit comfortably for the machine to detect the possible illness. In this regard, the researcher had no choice but to respect research ethics whereby the participant has a right not to divulge the information if he or she feels uncomfortable to do so. These findings contradict the assumption shared by most westerners that people automatically abandon their traditional medicine when acquainted with western medicine (Schott & Henley, 1996, p. 17).

The findings on pluralistic tendencies in healing support what was argued by Gordon (2010, pp. 165-183) and Varga and Veale (1997, pp. 911-924) that people continue using traditional medicine despite biomedicine's hegemonic attempts to prohibit such practices. Hence the biomedical health centres even appear to be contradicting themselves as explained above whereby they would be adamant that they do not promote the use of traditional healing methods while on the other hand they allow adult patients to use their traditional medicine only after four hours after the biomedicine has been applied. Local people have confessed that they continue the use of traditional healing methods in their homes for spiritual healing, despite the prohibitions of biomedical practitioners, as they strongly believe in the existence of spiritually-caused illnesses. Traditional people have demonstrated their scepticism regarding foreign or biomedical healing methods for *abantu* illnesses. This confirms assertions by Perret (2010) and Wedel (2009, pp. 49-64).

In biomedical healing, experimentation and scientific evidence in the laboratory reveal the causes of illness and determine the healing method. With the so-called theological illnesses and metaphysical or traditional healing methods, enculturation ensures the continued belief in the causes and cures of illness over decades and indeed centuries, and this longevity of the practices is evidence enough to sustain their belief system. As one of the participants (a community member and a nurse) revealed that there are even laws which insist on the use of exclusive healing method (biomedical healing method) for children, biomedical healing method seems to be the most popular with governments with political

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

support related to tangible evidence of successes. As far as governments are concerned, there is no place for metaphysical healing methods as their evidence is intangible and undocumented. In this case, it is deemed intangible despite the extended period of continued use and the strong belief systems around illness, health and medicine - hence the categorisation into superior and inferior cultural medicine – ethnomedicine - begins to take shape.

Linking the Source of the Healing Power with the Continued Use of Traditional Healing Methods for *Ukuhanjwa*

There are specific factors which contribute towards the trust and belief that Southern Nguni people have in the power of traditional healing for *ukuhanjwa*. These factors include the person who has recommended the traditional healing method, the bitter and biting taste of the medicinal plants, the very strong smell of the plants and the level of faith and belief that the caregiver has in the methods used.

Researchers have come to a variety of conclusions in relation to the continued use of traditional healing methods. Some of these are refuted by the outcomes of this research amongst the Southern Nguni people of ORTDM. These conclusions include the opening hours of service for traditional healers, whose constant availability is thought to contribute to the loyalty people have for them (Friend-du Preez, Cameron & Griffiths, 2009, p. 345; Ross, 2010, p. 47; Truter, 2007, p. 60). Research for this study suggests that the loyalty people have towards traditional healers and methods has little to do with opening hours, and far more to do with the power of belief, enculturation and conviction about the realities of the spiritual world. Biomedical centres are seen as simply not equipped to deal with the full nature of sickness. One female participant elaborated more on this and indicated that when the hospitals are unable to heal the child and not knowing what to do, they begin to ask questions about the medicine one has given to the child before opting for the biomedical health centre. After realising that they are unable to heal the ill child, the participant stated that they immediately ask the caregivers or mothers to take the child to the place where they started before opting for the biomedical health centre. To the Southern Nguni people, that is a sign indicating that they are also stuck but will not admit to having given up. The participant acknowledged that community members cannot speak openly about their indigenous healing methods to the biomedical health workers and practitioners because they end up being insulted about using their indigenous medicines. Those are some of the reasons which cause them to use indigenous medicine secretly.

Discouraging traditional people to use their own indigenous healing methods in favour of biomedical healing methods, the biomedical health workers have given traditional people the impression that biomedical healing methods are superior and to traditional people, superior means having absolute success in every performance. They are not familiar with the idea that even biomedical healing methods may fail, and that many treatments do not work perfectly the first time. When something that is touted as superior fails to work the first time, they do not think of requesting an adjustment to the dosage, or a different approach. They conclude that the disease is beyond the abilities of the biomedical centre – in effect, invisible to practitioners. It would appear that the idea of trying, adjusting, co-operating with the doctor to observe changes and improvements is not something that has been explained to people. Perhaps it is a mark of the hegemonic tendency of the biomedical health system that in their determination to win the battle against traditional practices, they too assume that their methods will work perfectly and do not consider that healing is sometimes a process, requiring trial and error in some cases.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Friend-du Preez, Cameron and Griffiths (2009, p. 345), Truter (2007, p. 60) and Ross (2010, p. 47)) have also mentioned the issue of staff and drug shortages in biomedical health centres as possible motivations for traditional people to continue with their pluralistic tendencies in healing. The outcomes of this research have indicated that neither shortages of drugs nor of staff are of chief significance, but what is of most significance is the Southern Nguni belief in the invisibility of cultural illnesses to biomedicine.

This brings the discussion back to the link between the perceived source of the healing power and the continued use of traditional healing methods for *ukuhanjwa*. The continued use of these methods has everything to do with the level of belief and faith that caregivers have in the efficacy of these healing methods. Another female participant admitted that Southern Nguni people take the victim of *ukuhanjwa* to the hospital so that they are not viewed as deviants and not because they believe and trust biomedical help. Southern Nguni people believe in indigenous medicine in healing this illness. According to this participant, even the biomedical nurses would go out and pick some leaves to make traditional medicine and apply it to the victim cautioning community members not to tell the doctor that the child has been given traditional medicine – those are nurses who have knowledge about this illness. It is important to note however that belief and faith alone would not have been enough to convince traditional people to continue with the use of traditional medicines. There has to have been evidence of healing with these methods. Thus the faith and loyalty that people show towards the traditional methods of healing for *ukuhanjwa* are backed up by a success rate that cannot be disputed – when it is caught early enough. The fear and alarm that this disease engenders is related to the fact that when noticed and acted upon too late, it invariably leads to death. That is why traditional Southern Nguni mothers always check the anus of their children when a child exhibits unusual sleepiness. Their belief and faith in the efficacy of the healing methods for *ukuhanjwa* is enhanced when the person who recommends the cure – or who recommended it to one's mother or other relative - has a history of positive cures, is trusted and a well-known diviner with an excellent track record.

The spiritual power contained in the plant, as ordained by God and revealed initially by the ancestors, also play their part in making these methods the trusted and apparently successful method that they are. So too does the bitter and biting taste of the plants. These are felt to owe their power to sympathetic magic by which body openings are shrunk back to their original size. Along with taste, the smell is believed to chase away the familiars or evil spirits until the victim has completely recovered as a result of sympathetic magic. The belief is that if the smell is strong enough to affect human beings, then sympathetic magic would make it possible for the familiars to be strongly affected by the smell too.

Ritual timing linked to the application of the healing method for it to work successfully to prevent escalation of the illness to the worst condition. It has always been emphasised that the indigenous healing method should be administered immediately after the illness is recognised and without any further delays for it to work successfully in the case of mild *ukuhanjwa*.

Ritual space is also linked to the maintenance of the healing power in the traditional medicinal plants. Some of the traditional medicinal plants for healing *ukuhanjwa* illness are believed to lose spiritual healing power should they be planted within the home premises – hence they are left to grow in the open fields. The belief is that in the case where these plants are planted within the home premises, familiars cause them to lose spiritual power by walking over them on their way to the victim for spiritual penetration. The same applies to the significance of the ritual space linked to privacy needed by the male owner of the familiar (*ichanti* snake) with the victim for efficacy of the healing method for *ukuhanjwa* illness. Hence no participant has divulged any information relating to the traditional healing method used by the male owner of the familiar – *ichanti*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The above sources of power have contributed mostly to the strong faith that the Southern Nguni have in the healing methods of *ukuhanjwa*. They would not have attempted to use them if the above sources were lacking. But from an outsider's point of view, one could conclude that continued use of these methods may be attributed to their efficacy and not merely to belief and faith.

SOLUTIONS AND RECOMMENDATIONS

The indications are that the Southern Nguni people of the ORTDM are deeply committed to their indigenous practices regarding *ukuhanjwa,* and are going to continue to consult with their diviners, use their preferred healing methods, and transfer their information on the illness from the present generation to the next.

The biomedical health system needs to pay attention to the interpretations traditional people give to their illness experiences for them to fully understand the point of view of the other person so that points of commonality and divergence can be identified and assessed. They may still experience difficulty in understanding and interpreting healing from a spiritual point of view, thus requiring the input of a traditionalist who would have to indicate if the illness requires spiritual healing, and whether it has been achieved or not. The two systems cannot be compared in a biological way as has been proposed.

Three approaches are used for the illness; the traditional, the biomedical, and the combination of both. This implies that there are three directions which could be taken. The first direction could be that of the biomedical health system being successfully convinced that *abantu* or cultural illnesses do exist and are real – something which might lead to the establishment of an integrated medical system – the Asian model (Baer, 2011, p. 417). The second direction could be that of traditionalists being convinced by the biomedical health system that *abantu* or cultural illnesses exist mostly in their imagination and have no basis in reality – a reality defined by westerners. This may lead to the exclusive medical system (only biomedical system) – the American model (Baer, 2011, p. 417; Levine, 2012, p. 66). The American position often reflects a narrow worldview where everything not familiar is, according to Levine (2012), "dangerous and irrational" (p. 66). The third direction could be that of maintaining the status quo of the inclusive medical system – the British model (Baer, 2011, p. 417) – a tolerant, yet condescending medical system, whereby the South African government extends tolerance to the traditional health system while preventing it from operating at the same level as the biomedical health system. It implies that traditional health care will remain in its lower status and the hegemonic tendency of the biomedical health system will continue. Should this be the case, there are nevertheless certain things listed below that biomedical practitioners can be made aware of, to inform their interactions with patients and improve their possibilities for successful healing.

- In the case of *ukuhanjwa*, despite its severity and possible fatal consequences, no preventative measures are used, despite their availability. This is because the illness is known to occur infrequently, and people can find little motivation to justify the continued alertness and vigilance that such preventative measures would imply. According to the Southern Nguni understanding, once the application of medicines for *ukuhanjwa* is discontinued, the smell is absent and the person is automatically vulnerable and open to attack. And yet no measures are taken to prevent this.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- The lack of a systematic way of preventing the illness has implications for health care in the broader sense. One might assume that as long as Southern Nguni people do not notice a threat to their lives, they find little reason to break with habit. For *ukuhanjwa*, sustained action over years would be required to be assured of its prevention, and people are unwilling to take this route. Similarly, they quickly fall away from any long-term practice adopted for prevention. Thus any medical protocol requiring consistency and patience is apt to fail, particularly if it is supposed to be adopted indefinitely with no likelihood of ever stopping. The same way of thinking seems to be evident in the reported resistance to using condoms for the prevention of HIV/AIDS and unwanted pregnancies. Recent reports on TV indicate that instead of increasing, the rate of condom usage in South Africa is dropping, with implications for HIV/AIDS and pregnancy statistics. A lack of consistency also extends to the treatment for HIV/AIDS; getting people to adhere to a long-term treatment regimen is extremely difficult. It has also been observed in the case of tuberculosis (TB), where patients are inconsistent in taking medications, and drop their use as soon as they begin to feel better, despite warnings not to do so.
- Thus, Southern Nguni people have demonstrated a preference for once-off measures, rather than long-term, consistent measures especially preventive measures. Prevention measure for polio, for instance, is dealt with in a more once-off manner, and there would be no requirement for consistency over the long term. Some people even state that there is no preventative measure for *ukuhanjwa*, whereas others confess that there is, it is just that people do not wish to engage with preventative measures. Hence with *ukuhanjwa*, the practice is to respond swiftly and thoroughly, but only for the short term. People do not underestimate the danger; in their view, they believe they are open to attack at any time; the problem is that people lack the ability to maintain consistency for extended period of time in the field of illness, health and medicine. Knowing these facts might enable biomedical practitioners to fully grasp what they are up against – it is not just forgetfulness, but a deep-rooted preference for dramatic, short term measures and a lack of impetus for consistent measures for prevention.
- Ritual and symbolism play a big part in Southern Nguni healing; indeed, an element of ritual is present in biomedical healing, too, and forms a part of all people's experience of healing. For Southern Nguni, there is a ritual time and ritual place for healing to take place and the practices for these things are accepted as uniform, with little deviation ever experienced. The usual experience is that rituals in the biomedical world are completely alien to traditional people. For example, naturalistic illnesses are accompanied by pain and spiritual illnesses are often not accompanied by pain; this is a ritualistic understanding which clashes with the biomedical understanding. Once pain is absent, people begin to seek spiritual solutions and this takes them in a different direction to the way of biomedical healing. Regarding the ritualisation of time and place, and the manner in which medicines should be administered, clashes arise between traditional beliefs and biomedical practices. The issue of symbolism and ritualisation in the biomedical health system is disguised and watered down, but it is present. People have come up with different forms of rituals and symbols in relation to illness and health that calm their nerves and help establish a sense of control. An example can be made of people who believe in the power of injections rather than taking pills. Their belief would be that they will be healed more quickly with an injection. To them the needle has symbolic value; they believe it will bring about a speedier recovery than pills, which have to go through the whole process of being digested before they begin to work, and this process is un-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

seen, time consuming and mysterious. We also see an element of ritual when people are instructed to take medication at certain times of the day and in a certain sequence. This ritual means a lot to the doctors, but is not understood by the traditional patients. For example a patient could be told to take a pill three times a day – in the morning, noon and evening. But when a patient realises that she has forgotten the morning pill, she may decide to take two pills at noon to cover for the morning pill she forgot, without questioning the implications of the timing. For as much as there is a ritual time and place in traditional medicine there is also a ritual time and place in biomedical healing for them to work properly, and these ritual performances may clash, leading to a lot of mis-conceptualisations.

- Ritual actions are apparent too. When a person takes the traditional medicinal plants for *ukuhanjwa*, the face tightens up as a result of the taste, and this is an important factor, indicating that the bodily parts will do the same when the medicine begins to work (sympathetic magic). The same principle applies to the oozing of the black liquid from the wound of the participant who believed that this oozing symbolised the actual healing, the release of poisons, after the application of indigenous medicine prepared from the sweet potato leaves, which had been put into her by witches. The participants' interpretation was confirmed by complete healing that was achieved after a relatively short time. As a result of these experiences with traditional medicines, people could very well expect similarly dramatic and visible evidence of healing when they visit biomedical health centres. They are accustomed to strong tastes, visible evidence and a degree of discomfort when experiencing healing. Slow, steady, consistent and unseen processes negatively impact on the trust of traditional people on the efficacy of biomedical healing methods. In addition, traditional people do not attach a high significance to the ritual timing stipulated by biomedical health centres concerning the application of healing methods. This could be because they are too accustomed to their own ritual timing and spaces, to the extent that they may even apply their own traditional procedures when using biomedicine as an expression of cultural modernity (using modern medicines in a traditional manner, rendering them less effective than planned).
- Southern Nguni people link the source of the healing power to the spiritual power of God in the creation of the bitter, sharp taste of the medicinal plants as well as God's creation of the very strong smell of the medicinal plants used for healing *ukuhanjwa*. The taste and smell indicate God's spiritual power in the features of these plants. Most biomedical physicians, though they may have a personal belief in God, do not ascribe the healing power in plants to God's power and as indicated by Davis-Floyd (1990, p. 277) their view of the human body is mechanistic, viewing the human body as a machine that could be taken apart and put back together to ensure proper functioning.
- There is also the issue of belief in healing methods. To biomedical physicians, the trust they have is in the chemical composition of the medicine and the concept of the source of that strength has little meaning for them. Spirituality has not been significant within the biomedical health sector unlike in the traditional health sector – hence the clash of cultures when it comes to illness, health and medicine. If doctors placed more emphasis on gaining the trust of the patient, that trust would become for the patient the source of the healing power. How people speak about the doctor (his or her reputation), and how he or she handles people as well as his or her methods of healing have to do with the spiritual healing of the patient, and this impacts positively or negatively on the physical healing of the patient. They all contribute to the trust the patient places in the physician.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

> Once trust is placed in a physician, much power will be ascribed to his medicines and methods, and his healing methods will be believed to work. Thus relationship is everything; it establishes a foundation of loyalty and devotion, which will have great implications for how patients adhere to protocols. Such practices could in turn minimise misunderstandings and help establish a common ground for social change and development which would not leave one culture feeling inferior and another, superior.

FUTURE RESEARCH DIRECTIONS

Issues discussed in this chapter are some of the knowledge gaps which need to be filled in order for biomedical practitioners to understand their patients, and perhaps adapt their own practises so as to enjoy greater success in bringing about the health and social change they desire. It could be a matter of finding areas of convergence and divergence between current practice and some of the ideas proposed, so that social change and development may be brought about in a harmonious way. The chapter has certainly not addressed all the issues surrounding illness, health and medicine that are relevant to harmonious social development as the focus was delimited to one *abantu* or cultural illness which has kept the findings and conclusion fairly narrow. A call is extended to other researchers to conduct further research in areas of concern regarding the practice of health, healings and medicine as an attempt to minimise misunderstandings and help establish a common ground for social development which would not leave one culture feeling inferior and another, superior.

CONCLUSION

Belief in a spiritual world and its manifestations in health and illness are generally assumed, by those who hold political and economic power, to spring from less well-developed cultures. The manner in which people conceptualise illnesses and illness causations determined and still continues to determine how they deal with illness and disease. The various approaches are culturally embedded and have less to do with history and development than with enculturation resulting in resistance traditionalism as a manner in which people hold strongly to traditional healing methods despite living in an era of globalisation and rapid change. Traditionalists who still believe in the efficacy of traditional healing methods are found in urban and rural areas, amongst the poor and the elite people and are literate and illiterate – hence the pluralistic tendencies in help-seeking amongst all these categories of people.

The different responses of different cultural groups to issues of health could be what triggered the hegemonic tendencies in the biomedical health system, which is one amongst several cultural healing methods. Its hegemonic tendency is clear, and empowered by the support it receives from governments the world over (Baer, 2011, p. 417). It tends to label the dominated forms of healing as inferior as part of a strategy to effect complete domination and bring about social change. The forces of globalisation on illness, health and medicine, social and political processes are often cited as reasons for the hegemonic tendency of the biomedical health system.

The Southern Nguni contextual meaning of invisibility of *ukuhanjwa* to the biomedical health system, and their conviction that it is a spiritual attack leads them to insist on traditional healing methods for the illness. Where biomedicine is sought, it is used as part of a pluralistic healing approach. The strong belief

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in the efficacy of traditional healing methods for cultural illnesses such as *ukuhanjwa* appears to be as a result of the perceived source of their healing power. This belief and faith in the source of the healing power enables traditional healing methods for this illness to withstand every attempt by biomedicine to undermine and counteract it.

Power relations and discourses existing in the field of illnesses and medicine are usually prompted by the language factor as well as the fact that no cultural group would submit to being made to feel inferior. Hence the traditional people tend to resist allowing extinction of their cultural practices by being placed in the past as history. Occasional failures of biomedical health centres to heal some of the illnesses have created loopholes which are used as reference by traditional people to promote their traditional/ indigenous healing methods. Their interpretation is that biomedical health centres fail to heal such illnesses only because they are invisible to them and are cultural/*abantu* illnesses. All these are attempts by the traditional people to resist the hegemonic tendencies of the biomedical health system. On the other hand, the biomedical health system would insist that their hegemonic tendencies are for the good cause, meaning that of enhancing evolution of culture as well as bringing about social change for the betterment of human health conditions.

Absence of pain to the Southern Nguni people means illness resulting from acts of witchcraft. They generally cannot imagine any other illness which could be without pain other than cultural or *abantu* illness. In the study, the Southern Nguni people gave examples of illnesses they experienced without pain including *ukuhanjwa*. They could even differentiate between different types of cries whereby their contextual meaning of a weary cry was tiredness and the contextual meaning of a loud cry was pain. Hence they linked the weary cry of victims of *ukuhanjwa* with tiredness and not pain.

Although the South African government has lifted laws of apartheid against traditional healing, it seems as if the act is just on paper but not operational according to the desires of the traditional people who would like to see operationalisation of integrated health system encompassing both the traditional and biomedical health system. What seems to be working for now is the inclusive and tolerant health system whereby the biomedical health system remains favoured by the government and recognised as the official and authorised health system. On the other hand, traditional health system is openly used by the traditionalists but the system is still demeaning medical system to the traditionalists as the system keeps the traditional health care in its lower status. The manner in which the Southern Nguni people responded indicated that they prefer the integrated medical system as currently they are still using their traditional or indigenous healing methods secretly. This simply means that, although it is well known that Southern Nguni people use traditional or indigenous healing methods for some of their illnesses, they do not openly admit to such practice to the biomedical health workers and practitioners.

REFERENCES

Baer, R. D., Weller, S. C., De Alba Garcia, J. G., Glazer, M., Trotter, R., Pachter, L., & Klein, R. L. (2003). A cross-cultural approach to the study of the folk illness nervios. *Culture, Medicine and Psychiatry*, *27*(3), 315–337. doi:10.1023/A:1025351231862 PMID:14510097

Barker, C. (2003). *Cultural studies: Theory and practice*. London: Sage.

Bhana, K. (1986). Indian indigenous healers. *South African Medical Journal*, *70*, 221–223. PMID:3738660

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Davis-Floyd, R. E. (1990). Ritual in the hospital: Giving birth the American way. In P. Whitten & D. E. K. Hunter (Eds.), *Anthropology: Contemporary perspectives* (pp. 275–285). London, UK: Longman Higher Education.

Dayes, J. E. (2014). *A pluralistic approach to medically unexplained symptoms*. (Unpublished Doctoral thesis). City University London, London, UK.

Douglas, M. (2010). *Purity and danger*. London: Routledge.

Eastman, M. (2011). Rainbow healing: Traditional healers and healing in South Africa. In T. W. Bennett (Ed.), *Traditional African religions in South African law* (pp. 184–199). Claremont, South Africa: UCT Press.

Farrand, D. (1984). Is a combined western and traditional health service for black patients desirable? *South African Medical Journal*, *66*, 779–780. PMID:6495131

Foster, G. M. (2009). Disease etiologies in non-western medical systems. *American Anthropologist*, *78*(4), 110–117.

Friend-du Preez, N., Cameron, N., & Griffiths, P. (2009). *Stuips, spuits* and prophet ropes: The treatment of *abantu* childhood illnesses in urban South Africa. *Social Science & Medicine*, *68*(2), 343–351. doi:10.1016/j.socscimed.2008.10.027 PMID:19028413

Good, B. J. (2001). *Medicine, rationality and experience: An anthropological perspective*. Cambridge, UK: Cambridge University Press.

Gordon, D. (2010). A sword of empire? Medicine and colonialism in King William's Town, Xhosaland, 1586-1891. *African Studies*, *60*(2), 165–183. doi:10.1080/00020180120100294

Hu, B., Du, Q., Shen, K., & Xu, L. (2012). Principles and scientific basis of traditional Chinese medicine in cancer treatment. *Journal of Bioanalysis & Biomedicine*: *An Open Access Journal*, *S6*(5), 1–6.

Kleinman, A. M. (2010). Medicine's symbolic reality: On a central problem in the philosophy of medicine. In B. J. Good, M. M. J. Fischer, S. S. Willem, & M. D. Good (Eds.), *A reader in medical anthropology: Theoretical trajectories, emergent realities* (pp. 85–90). West Sussex, UK: Wiley-Blackwell.

Leclerc-Madlala, S. (2002). On the virgin cleansing myth: Gendered bodies, AIDS and ethnomedicine. *African Journal of AIDS Research*, *1*(2), 87–95. doi:10.2989/16085906.2002.9626548 PMID:25871812

Levine, S. (2012). Testing knowledge: Legitimacy, healing and medicine in South Africa. In S. Levine (Ed.), *Medicine and the politics of knowledge* (pp. 55–78). Cape Town, South Africa: HSRC Press.

Liebeck, H., & Pollard, E. (1995). *The Oxford large print dictionary*. Oxford, UK: Oxford University Press.

Lupton, D. (2012). *Medicine as culture: Illness, disease and the body*. London: Sage.

Mofokeng, M. A. (2003). *Beliefs and practices of Sotho antenatal women*. (Unpublished Master's dissertation). University of South Africa, Pretoria, South Africa.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ngqila, K. H. (2002). *Beliefs in spiritual causation of infant illnesses and their mortality in the Tafeni Location of Ngqeleni District (Eastern Cape)*. (Unpublished Master's dissertation). Walter Sisulu University, Mthatha, South Africa.

Perret, D. (2010). *The science of spiritual healing: Preparing for deep changes*. Paris: Books on Demand.

Petrus, T. S., & Bogopa, D. L. (2007). Natural and supernatural: Intersections between the spiritual and natural worlds in African witchcraft and healing with reference to South Africa. *Indo-Pacific Journal of Phenomenology*, *7*(1), 1–10. doi:10.1080/20797222.2007.11433943

Ross, E. (2010). Inaugural lecture: African spirituality, ethics and traditional healing – implications for indigenous South African social work education and practice. *South African Journal of BL*, *3*(1), 44–51.

Schott, J., & Henley, A. (1996). *Culture, religion and childbearing in a multiracial society: A handbook for health professionals*. Oxford, UK: Butterworth-Heinemann.

Shoko, T. (2011). Shona traditional religion and medical practices: Methodological approaches to religious phenomena. *Africa Development. Afrique et Developpement*, *36*(2), 277–292.

Thinwa, J. (2004). *Indigenous healing practices and their effect on TB and HIV/TB patients' utilization and compliance with anti-TB medication*. Retrieved November 28, 2012, from http://digitalcollections.sit.edu/isp_collection/498

Truter, I. (2007). African traditional healers: Cultural and religious beliefs intertwined in a holistic way. *SA Pharmaceutical Journal*, 56-60.

Urbasch, M. (2002). African traditional healing systems: Representations and restitutions. In P. Guillaume & A. Shepperson (Eds.), Bodies and politics: Healing rituals in the democratic South Africa (pp. 1-76). Braamfontein, South Africa: French Institute of South Africa 2.

Varga, C. A., & Veale, D. J. H. (1997). Isihlambezo: Utilization of patterns and potential health effects of pregnancy-related traditional herbal medicine. *Social Science & Medicine*, *44*(7), 911–924. doi:10.1016/S0277-9536(96)00104-9 PMID:9089914

Vaughn, L. M., Jacquez, F., & Baker, R. C. (2009). Cultural health attributions, beliefs, and practices: Effects on healthcare and medical education. *The Open Medical Education Journal*, *2*, 64–74.

Wedel, J. (2009). Bridging the gap between western and indigenous medicine in Eastern Nicaragua. *Anthropological Notebooks*, *15*(1), 49–64.

Whiting, B. B. (1977). Paiute sorcery: Sickness and social control. In D. Landy (Ed.), *Culture, disease and healing: Studies in medical anthropology* (pp. 210–218). New York: Macmillan Publishing Co., Inc.

KEY TERMS AND DEFINITIONS

Abantu **Illnesses:** Spiritually caused illnesses which can only be healed using traditional healing methods.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cultural Modernity: Anything which is considered to be modern but still has traces of the culture of a particular group of people. This could be clothes, type of education, types of behaviour, etc.

Indigenised: Anything that is considered to be modern and is used in ways not foreseen by the manufacturer in order to suit the needs of the group of people who use it.

Indigenous: Anything that is endemic to a place or a group who use it in ways it was originally intended to be used.

Invisibility: The Southern Nguni contextual meaning of invisibility of *abantu* illnesses is means the inability of the outsider to link the symptoms of *abantu* illnesses to their causes as understood by the Southern Nguni people in order to be able to prescribe the correct healing method.

Pluralistic Healing Method: The use of traditional and biomedical healing approaches simultaneously or use of traditional healing approach in the beginning and at a later stage use of biomedical healing approach as a last resort or use of biomedical healing approach in the beginning and at a later stage use of traditional healing approach as a last resort.

Ritual Purification: Any method which is used to heal a spiritually-caused illness or cleanse a polluted person so as to prevent illness or misfortune. This cannot be equated with physical cleanliness.

Ritual: Religion in action whereby an act is done in a particular situation in a set precise manner which is the same each time it is performed for the relief of anxiety. The term has also been circularised whereby a ritual becomes any act that is regularly repeated in a set precise manner.

Traditional Medicine: Medicine that is based on theories, beliefs and experiences indigenous to different cultures used in the maintenance of health as well as in the prevention, improvement and treatment of physical and spiritual illnesses. Traditional medicine could be both indigenous and indigenised.

Traditional People: People who still strongly believe in traditional practices irrespective of their level of education (literate or illiterate), location (urban or rural), or status (rich or poor). These traditional practices could be purely indigenous, or indigenised.

This research was previously published in the Handbook of Research on Theoretical Perspectives on Indigenous Knowledge Systems in Developing Countries edited by Patrick Ngulube; pages 375-394; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 52
Kinect-Based Limb Rehabilitation Methods

Yongji Yang
Changchun University, China

Zhiguo Xiao
Changchun University, China

Furen Jiang
Changchun University, China

ABSTRACT

Within the context of health informatics, this article discusses how real-time information of human skeleton movements can be conveniently captured via the use of Kinect-based deep sensor. It highlights an effective action recognition method, using Unity3D to form virtual models of characters and scenes. Kinect somatosensory cameras can now identify one's motion data and provide feedback on the virtual role model to complete the avatar drive. This is to achieve real-time simulation modeling as well as drive operation instructions based on the design and identification verification to complement limb rehabilitation. The hardware of such a health informatics-related system is simple, inexpensive, and highly meaningful in term of augmenting user experience. In conclusion, the suggested approach can contribute significantly towards aiding physical rehabilitation therapy

1. INTRODUCTION

Traditionally, rehabilitation medicine entails the deployment of either man-assisted and/or robot-aided rehabilitation methods to enhance the physical functioning of individuals affected by illnesses or injuries with the goal of minimizing any preventable activity-related impairment. With a growing shortage of medical resources, the rising cost of specialty medical equipment, and its high maintenance costs, a call for new inexpensive means of rehabilitation medicine is highly advocated. One of these innovations is the design of intelligent human-computer interaction (HCI) within the emerging health informatics context.

DOI: 10.4018/978-1-7998-1204-3.ch052

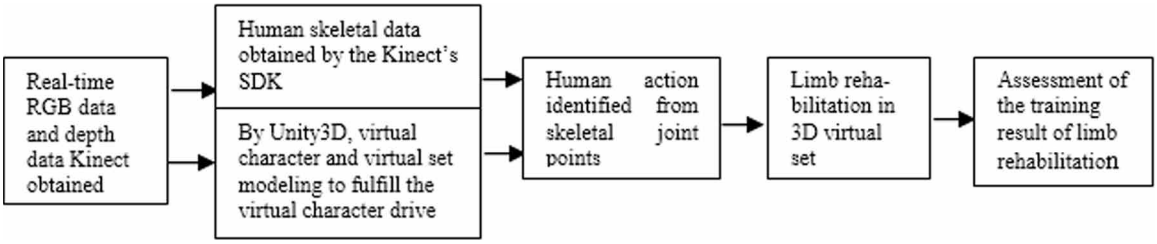
Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

HCI refers to the transmission of one's ideas to the computer through kinesics (such as gestures, postures, facial expressions, and more) as well as the computer recognition of human action and behavioral feedback. Using HCI to replace traditional rehabilitation methods can effectively resolve the cost and medical resource shortage challenges as it provides the patients with a convenient, inexpensive and effective rehabilitation system. Today, HCI is trending in current rehabilitation research. Yet, temporal variability and the complexity of human action is a key challenge in promoting action recognition. At present, the method of motion parameters collected by wearable sensors, gyroscopes, and accelerometers for the recognition of human body movement is accurate and highly effective (Maekawa, Yanagisawa, & Kishino, 2010; Cao, Cai, & Cheng, 2010), but the wearable sensor reduces the wearer's comfort. In term of visual recognition method, the current study draws on 2D visualization, with the effect often being highly influenced by environmental background, illumination, and occlusion factors.

With the emergence of inexpensive Kinect somatosensory cameras, new opportunities have arisen for action recognition researchers. With the aid of skeletal tracking technology to gather real-time data, these researchers can now collect non-contact 3D human skeleton information, which is hardly influenced by the background illumination. There is also direct use of data based on the key points of Kinect motion recognition method, a method to set the rule to identify joint action between the use of position angle, height and other mobile velocity, acceleration, and the like to define thresholds. Xie Liang et al. (2013), for example, obtained good recognition results by determining the specific posture based on the Euclidean distance and angle between the joints. Such a method is now widely applied due to its ease of use, but it displays low extensibility and a limited recognition in newly defined domains after each new action to be inserted. Another approach is to extract features, then select the appropriate classification for recognition. Raptis et al. (2011), for example, used Kinect to get bone information, transforming it to human angle features from the perspective of the human skeletal structure, which can achieve the classification of dance gesture recognition and the accuracy can be upward of 91.9%. Lai et al. (2012) transform the joints features into 2D features, and the nearest neighbor classifier is used to recognize the hand gesture, which can make the recognition rate to reach 92.25% under the condition of limiting the speed of movement. In this work, we propose using Kinect recognizing skeleton key information, by feature selection, focusing on the temporal and special problems of motion recognition to achieve flexibility, robustness, and high instantaneity. Such an approach can then be used for limb rehabilitation action identification; moreover, this same approach, by binding human skeleton information identified by Unity to the 3D animated character, allows users to practice limb rehabilitation training in a virtual 3D scene as schematically depicted in Figure 1.

Figure 1. Proposed Kinect-based Limb Rehabilitation Method



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. BACKGROUND

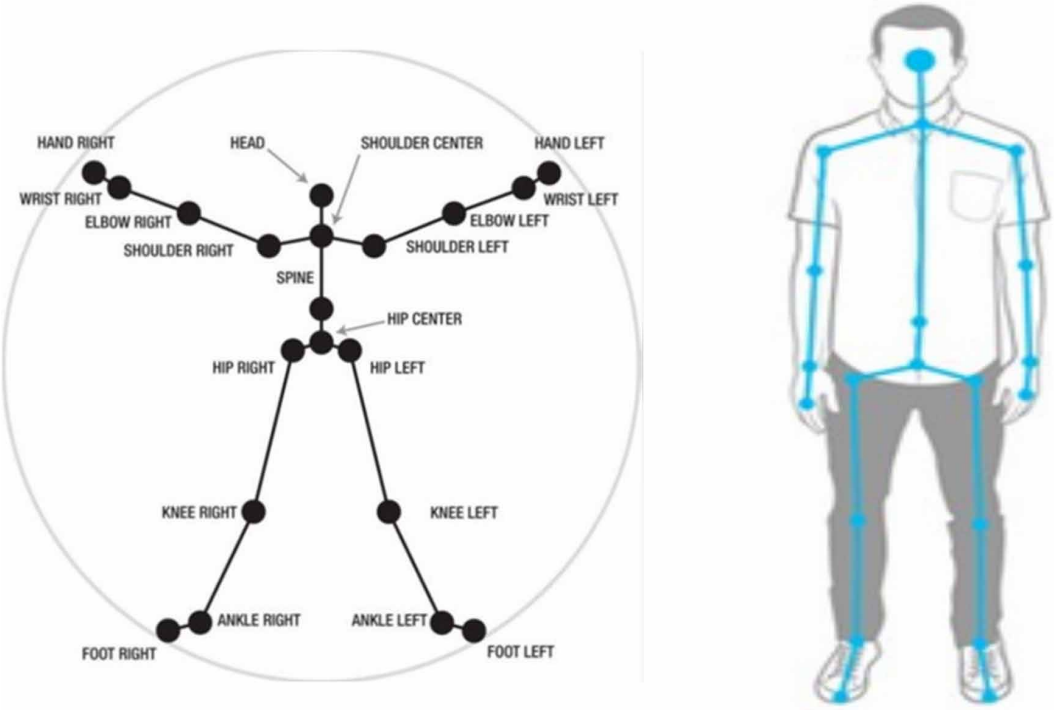
Kinect deep cameras can capture the colored image and deep image of the surrounding scene. In (Lin, Hsieh, & Lee, 2013), the data source used was real-time 3D information obtained from 20 individual body bone joints provided by the Kinect SDK application programming interface (API). Together, the captured information can fully represent a complete and effective body construction in a variety of postures. As shown in Figure 2, the distribution of human skeleton joints captured by Kinect cameras is a depiction of those bone joints that are the posture traits of the body postures. Later, we will discuss more about the matching algorithms for identifying those postures involved in the feature point position below.

Microsoft bone-tracking technology to capture via Kinect for configuring the location of main human joint points and tracking the deep image may be divided into the following three steps (Shotton, Fitzgibbon, Cook, 2013):

Step 1: Foreground Segmentation

Initial stage comprises the use of image-processing technology to achieve human foreground segmentation, which is, analyzing each pixel according to distance to see whether or not the captured image is part of a human body. Accordingly, relevant images of the human body parts are extracted from the depth of the image as a whole through the respective edges according to distance relationship.

Figure 2. The human bone joint points captured by Kinect



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Step 2: Body Parts Recognition

This stage involves identifying the various parts of the body from the body contours of the image area. Essentially, the step uses hundreds of TB meter training data obtained from Random Forest model for classifying parts of the body. This model can give each pixel a category label and probability value(s) to be associated with the different body parts (such as the head, shoulder, and so on with the entire human body divided specifically into 32 different parts).

Step 3: Joint Points Recognition

The pixels which belong to the same part are combined into one joint point, and the joint location is determined by the Mean Shift local model identification method based on the Gauss kernel function; more specifically, the primary task is to locate the human bone joints accurately from the front and side angles.

Based on the above three steps, researchers can get the bone image data. The current Kinect visual approach can recognize and locate 20 joint points, with 3D information of each joint point. This paper presents a set of real-time contrasting experiments based on Kinect skeleton-based pose identification method from the health informatics perspective as proposed by T. Y. Lin et al. (2013). In this work, a series of experiments, under the same conditions, has been carried out with the method described above and the average time of recognition of each kind of action posture is computed. Compared with the recognition method of TYLin, the average recognition time of each kind of action posture in this method is shorter with the average time for recognition of each action posture recorded as 41.3ms, and the average recognition time of the method being 33.8ms. Altogether, the experimental results show that the proposed method has better real-time performance.

3. ALGORITHM DESCRIPTION

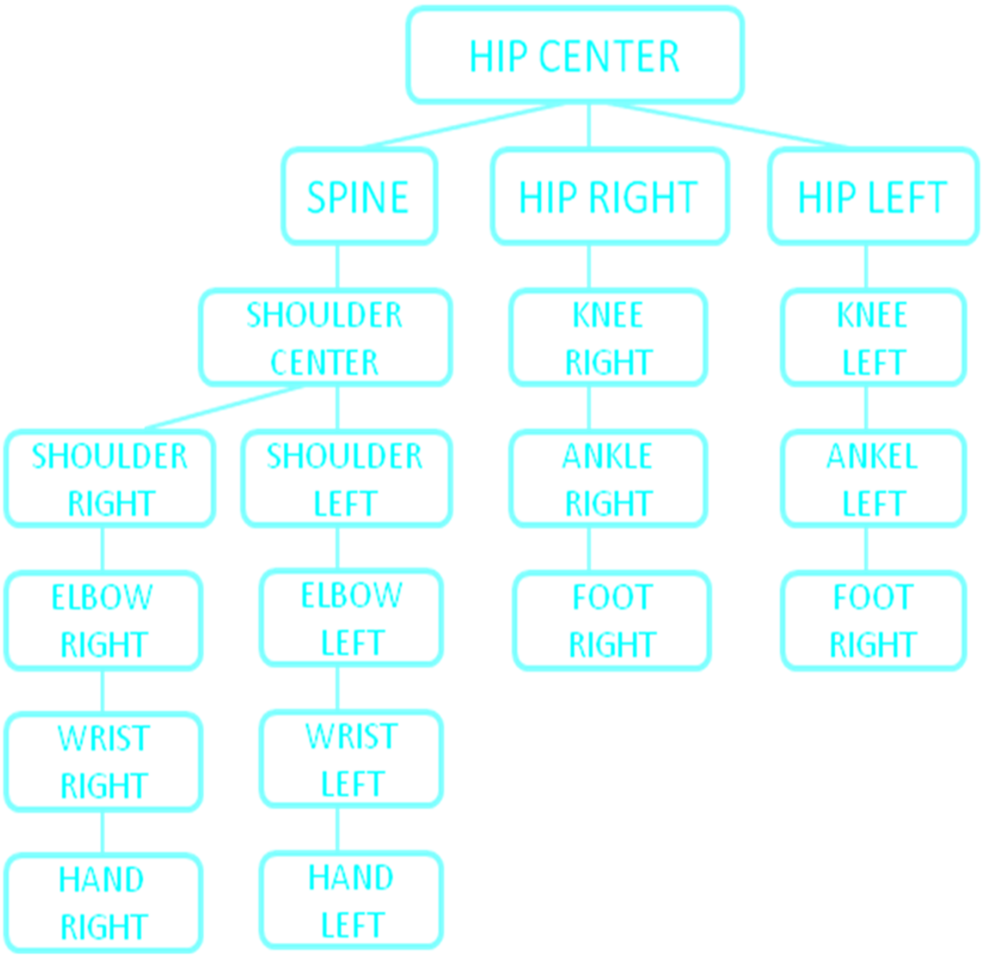
3.1 Skeletal Point Feature Extraction

In human action recognition, the expression of 3D skeletal point is commonly the three-dimensional position coordinates and the amount of rotation of the three axes. The 3D position coordinates are more sensitive in terms of users' size and height and often require additional normalization; moreover, the amount of rotation around the 3 axes is prone to universal lock (Gimble lock). In contrast, the rotation information of the skeletal point highlighted in this paper represents the skeletal point information in space. This is often expressed by quaternion (quaternion) which will effectively avoid the normalization and gimbal lock problems. To define the rotation information skeletal point, you must define the skeletal point coordinates. Here, the center point of the hip is considered the bone coordinate origin of the skeleton, which has no rotation angle. The adjacent bone points extend downward, stratified by about 20 bone points and are representative of the lower nodes. As shown in Figure 3, the adjacent skeletal points have been divided into "parent" nodes, which are the upper bone points versus the "child" nodes, which are the lower bone points.

For all skeletal points detected, we use a 4D vector (x, y, z, w) to describe the spatial information of the skeletal point where (x, y) refer to the skeletal point where the bone coordinate position is set corresponding to the skeletal point of the image location (x_m, y_m); thus, if the hip image position is (x_n, y_n), then $x = x_m - x_n$, $y = y_m - y_n$. Likewise, z is the skeleton point corresponding depth value, whereas w is for the skeletal point and its parent node vector forming an angle with the positive direction of the x-axis; so, if the parent node is skeletal point coordinates (x_1, y_1, z_1, w), then w by formula (1) is calculated as:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Hierarchical Diagram of Human Skeletal Point



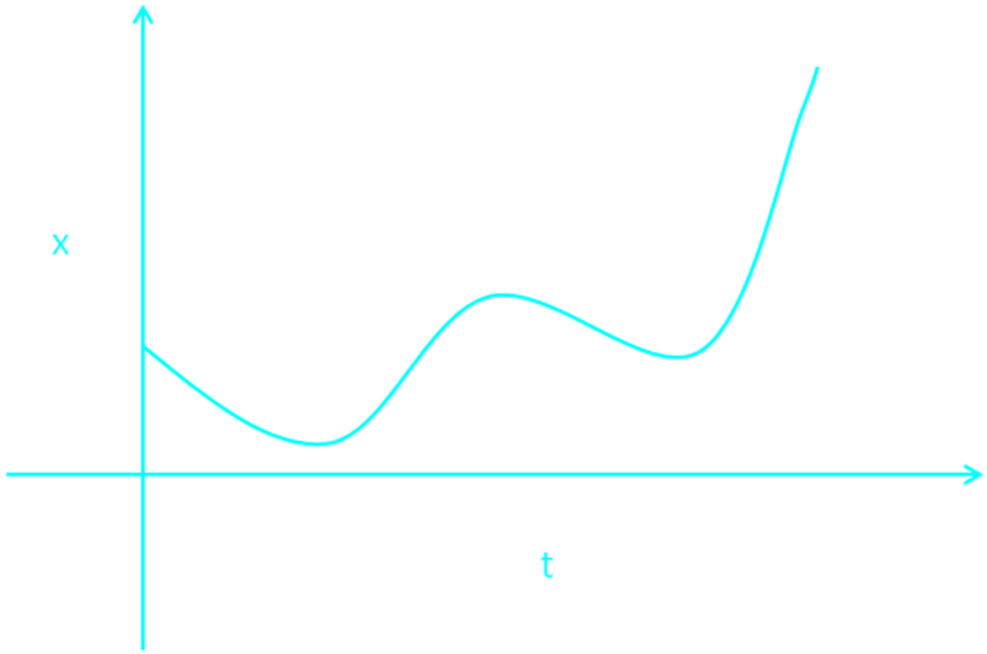
$$w = \arccos \frac{x}{\sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2}} \tag{1}$$

3.2. Weak Classifier Training

The most essential technology in using Kinect to carry out limb rehabilitation training in the health informatics context is to determine whether the body movements are correct and standardized. Hence, the dynamic recognition of human actions needs to be examined closely alongside the changes of skeleton data in a row. Experimental results show that certain action of a skeletal point may be represented as a weak classifier for physical rehabilitation recognition since computation of the parameter of any such skeletal point must follow certain rules. Figure 4 shows changes of x coordinate values when a skeletal point is doing an action. Although the speed of Kinect capturing images is up to 30 frames/s, the changes of the body movement are generally continuous that will not change within a short period of time, and therefore the performance operating characteristics can be fully displayed with 1/s intervals value evalua-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Graph of the x component of the skeletal point over time



tion. Our experiment uses 20s as a period; thus, such a weak classifier characteristic is a 20-dimensional vector. As this features the combination of time segment, it is called space-featured vectors. Given that a skeleton has 20 bone points, each bone point has four parameters, then we can possibly have 80 weak classifiers to fulfill the training.

For categorizing each of the weak classifier, we use SVM (Amami, AyedD, 2013; Yang, Kim, & Park, 2012) to train so as to derive the classification function of each weak classifier. The micro-steps involved are as follows:

1. Define the spatial and temporal characteristics vector set x. A space-time vector set comprises a temporal feature vector information for all skeletal point; that is, a space-time vector set contains 80 20-dimensional vectors, with a time span 20 s, i.e., $x = \{x_1, x_2, ..., x_{80}\}$, each vector xi represents a vector space-time coordinates of a bone point.
2. Take the *nth* currently-to-be-trained recognition action A as temporal and spatial motion vector set, then once again take the *nth* other various types of actions as space-time motion vector set.
3. Train each space-time vector SVM, you can get a total of 80 decision-making functions, thus identify every time and space vector.

After the above training, you will get access to the spatial and temporal characteristics of the vector (Ye, 2014) weak classifiers.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.3. Action Feature Set & Achieving a Strong Classifier

SVM classifier can make classifications of the newly input samples trained by each space-time vectors, but classification accuracy typically cannot meet the demand; hence, using AdaBoost algorithm, weak classifiers can be combined into a strong classifier which can greatly enhance recognition rate (Yim, & Lea, 2014; Chao, Xilu, & Su, 2016). The specific training process is detailed as below:

(1) Getting the training sample

For some action A, we obtain a total of 2n of a series of space-time sample set (x_1, y_1), (x_2, y_2) ... (x_{2n}, y_{2n}), and *nth* samples positive and *nth* samples negative, wherein $y_i \in \{0,1\}$, $y_i = 1$ shows a sample xi which belongs to the action A, negative sample.

(2) Circuit train all groups

$$w_i = 1 / 2n \quad i = 1,2, ..., 2n \tag{2}$$

1. initialized weights w_1, $x = w_x$ where w_1,x is initialized right of the first round of training weights.

For t = 1: T, where T represents a total training of T rounds, t is the current training round number.

2. weight renormalization

$$q_{t,i} = \frac{w_{t,i}}{\sum_{y=1}^{2n} w_{t,y}} \tag{3}$$

where, $q_{t,i}$ is the normalized weights.

3. Each temporal sample set includes 80 spatial and temporal feature vectors, set the *mth* eigenvector to identify, m = 1,2, ..., 80. Since having obtained the corresponding spatial and temporal characteristics of SVM classifier, it is possible to calculate the *mth* feature vectors recognition rate in this group of training samples, wherein the *mth* weighted error rate.

$$\varepsilon_{\int}^{m} = \sum_{i} \left(q_t, h\left(d\right)_{t,i} \right) \tag{4}$$

In formula (4), h (d) t,i is judging result of the *mth* characteristic in Sample i, if it is judged correctly, then $h(d)_{t,i}=0$, if judged wrongly, then $h(d)_{t,i}=1$. Repeat step iii, you can get all the spatial and temporal characteristics of the weighted error rate of *m*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. Select the lowest round of the weighted error vector of the spatial and temporal characteristics.

$$\varepsilon_t = \min\left(\varepsilon_{\int}^1, \varepsilon_{\int}^2, ...\varepsilon_{\int}^m\right) \tag{5}$$

5. Adjust the weight of each group according to vector results:

$$w_{t+1,i} = w_{t,i}\beta_t^{1-e_i} \tag{6}$$

In the above formula, $e_i = 0$ indicates that this group is a correct identification whereas $e_i = 0$ indicates that this group is misclassified, and $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$.

(3) Get a strong classifier

After training of T rounds, a total of T spatial and temporal characteristics obtained as the final classification features, the last strong classifier can then be calculated according to formula (7).

$$C(x) = \begin{cases} 1 & \sum_{t=1}^{T}\alpha_t h_t(x) \geq s\sum_{t=1}^{T}\alpha_t \\ 0 & other \end{cases} \tag{7}$$

Wherein, $h_t(x)$ is the judging result of the *tth* space-time vector of the input actions recorded which is regarded $h_t(x)=0$; s is less than coordinate 1. If it is closer to 1, the recognition-erroneous rate will be reduced; meanwhile, the recognition rate will gradually increase as well as the erroneous rate, and generally the most suitable s is $s = 1/2$.

3.4. The Body Posture Optimization

When Kinect interacts with humans, its infrared sensors need to track every part of the body. In order to achieve higher recognition accuracy, this method requires the users to face the equipment without being blocked. Once obscured, the Kinect system will fail to track the joint points accurately, leading to poor judgments on some parts of the human body. To avoid serious distortion, a good Kinect application not only requires equipment to have clear recognition effect not interrupted by any noise, but can still achieve high recognition accuracy under obscure and noisy environment. In light of this, the authors hereby propose a human bone reconstruction method based on reliability assessment, which can rebuild the defective skeleton pose by means of reliability evaluation and motion database. Several steps are involved: first, the reliability after getting the user's skeleton movement value is calculated and subject is placed as the weight, then similar posture in motion database is extracted. Following this, a natural position space in combination with principal component analysis to satisfy the kinematic constraints is built to finally synthetize natural accurate skeleton measurements. Experimental results show that, when obscured with different degrees, bone reconstruction method based on reliability assessment can still reconstruct movement skeleton of comparatively high quality.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Skeleton reconstruction method is composed mainly of skeleton acquisition, movement reliability assessment, and the building of natural posture space by the use of database structure, skeleton synthesis, and more. Among these, the movement reliability assessment is the most important. Reliability evaluation refers to the objective evaluation of the following three aspects, more specifically, the stability and reliability of the various parts of our body skeleton including:

(1) Behavior reliability
(2) Skeleton reliability
(3) Tracking reliability

- Behavior reliability

Behavior reliability refers to the correctness of the action behavior. When a user's body parts are blocked, the depth image captured by Kinect will mark as if it is impossible to get the accurate location of one's body parts. Due to the lack of the expected target features, positions identified by the method without reliability evaluation are very much unstable (Ji, & Duanyuan, 2016; Zhou, 2015). Therefore, the method should at first weigh the reliability of the user behavior.

Assuming that pi(f), pi(f+1), pi(f+2) are the three consecutive frames tracked body part in 3D space, so, the calculating formula for the displacement is

$$d_i(f) = P_i(f+1) - P_i(f) \tag{8}$$

$$d_i(f+1) = P_i(f+2) - P_i(f+1) \tag{9}$$

The Angle between two vectors can be calculated by the dot product:

$$\theta_i(f) = \begin{cases} \dfrac{d_i(f)d_i(f+1)}{\left|d_i(f)d_i(f+1)\right|} & \left|d_i(f)\right| > d_{\min} \cup \left|d_i(f+1)\right| > d_{\min} \\ 0 \end{cases} \tag{10}$$

The d_{min} is the minimum of the acceptable displacement, when body parts are basically stable, there will not be a big angle change.

Human behavior can be seen as a series of motion sequence. Here, behavior reliability may be defined as:

$$Rb_i(f) = 1 - \frac{\max\left(\min\left(\dfrac{\sum_{f=0}^{f_b}\theta_i(f)}{f_b}, \theta_r\right) - \theta_f, 0\right)}{\theta_r - \theta_f} \tag{11}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Among them, $Rb_i(f) \in [0,1]$, when f_b, the total number of motion frames is detected, θ_f is the smallest rotation angle, θ_r is the smallest unacceptable angle of each frame, repeated experiments show that, when the fb = 3 $\theta_f = 90°$; $\theta_r = 135°$, the effect is better.

- Skeleton reliability

Skeleton reliability refers to the correctness of the skeleton position. When Kinect captures body postures, the user's skeleton size is different from one user to another. Defining two positions for reference, one is T pose, and the other is the standard standing posture. According to the skeleton structure, each part can be connected to many other parts of the body. Assuming that the body parts I were connected to a St of other parts, then for each connection section of s, in the first frame differences f rate can be calculated as follows:

$$d_s(f) = \min\left(\left|l_s\left(f\right) - l_c\right| \Big/ l_c, 1\right) \tag{12}$$

l_c is the length of the related frame part, $l_s(f)$ is the current length of the first frame f s period of skeleton.

Motion reliability is the average period of different proportions all connected, and is defined as:

$$Rk_i(f) = 1 - \sum_{s=0}^{s_t} d_s(f) \Big/ s_t \tag{13}$$

In which, $Rk_i(f) \in [0,1]$.

- Tracking reliability

Tracking reliability refers to whether or not the equipment can trace the body part of the user, as defined by:

$$Rt_i = \begin{cases} 1.0 & \text{Be traced} \\ 0.0 & \text{Not traced} \end{cases} \tag{14}$$

Obtaining the reliability of behavior, skeleton and the state of tracking by the above assessment, the reliability of different body parts can be defined as:

$$R_t = \sum_{f=0}^{f_{total}} \omega(f) \times \min\left(Rb_i(f), R_i\right) \tag{15}$$

In which, $R_i \in [0,1]$, f_{total} is the total number of frames, $\omega(f)$ is f frame gaussian weighting, $Rb_t(f)$, $Rk_t(f)$, $Rt_t(f)$ respectively refers to reliability, motion reliability and tracking reliability. When the body part is not being tracked, it will extract the adjacent frames and reliable data for reconstruction, thereby improving the precision of the identification.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A natural position space needs to be built for human posture reconstruction, which needs a database also to be built. Database building can use the traditional optical system to capture the right sports movements. In addition to the improvement of the operating efficiency, it can also control the density of the sample position space. The movement database in this paper is from Cameron university (CMU) graphics laboratory. Motion data formats are: ASF - AMC, HTR, BVH, etc. The ASF-AMC data formats are applied here, which can further modify the motion data and bone model as part of additional contributions of the present study.

3.5. Action Identification and Analysis

As a starting point of discussion, we select here five representative body movements from those common actions of the body rehabilitation for experimental research, for example, the right arm lifting, upper limb external rotation, the right leg flat lifting, the right leg side lifting, and squatting. The strong classifier obtained from training is binary classification problems. In order to identify the five different actions, the authors collected 3000 bone data for each bone movement when collecting samples, took 30 consecutive frames of each movement of the bone data as action positive samples whereas other actions as negative samples of this action. Altogether, we now have 100 positive samples obtained for every movement, plus 400 negative samples. The conclusion is that the classifier can distinguish five different actions.

Here, our experimental testing samples are the skeletal human data collected from five persons of different height and size. 6,000 bone data are collected from a single action, which constitutes 200 sample sets, with a total of 1,000 action samples collected. During sample identification, based on the action feature set, sample data are divided by five strong classifiers, whose results are identified and released. If the output of two or more strong classifiers is 1, this means that the sample cannot be identified. The outputs and recognition accuracy of 1000 samples are shown in Table 1.

In order to improve recognition accuracy, the authors correspondingly combine the aforementioned five strong classifiers. A 5-bit binary numbers are used to represent actions. From left to right, they are right arm lifting, upper limb external rotation, right leg flat lifting, right leg side lifting, and squatting. The corresponding bit 1 indicates that the corresponding strong classifier identifies the action, for example, a binary number 01000 represents recognition of the action—the right arm lifting. Overcoming

Table 1. Recognition rate of the first five actions combined

Action	Output the right arm lifting	Output upper limb external rotation	Output the right leg flat lifting	Output the right leg side lifting	Output squatting	Recognition failure	Accuracy rate
the right arm lifting	175	15				10	87.5%
upper limb external rotation	10	172				18	86%
the right leg flat lifting			175	10		15	87.5%
the right leg side lifting			8	178		14	89%
squatting					156	44	78%

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the emergence of the first digit or more than five digits is 1 which will result in recognition failure, we propose action recognizing intensity of the strong classifier P, which holds strong classifier detection rate before P represents a strong classifier training; as such, the higher the detection rate, the greater P is. For example, the binary number 00110, $P_3 > P_4$ indicates the recognition of the action of the right leg flat lifting, otherwise it means recognition of the action of the right leg side lifting. Because the strong classifiers are a few good-weighted weak classifiers intensified, larger action recognition intensity P indicates the weak classifier corresponding to strong classifier identifying the actions or larger weights actions. So, the recognition accuracy of these strong classifiers compared to that intensity of P, is much higher. Experimental results show that the combination, compared with the previous one, has greatly improved action recognition rate. The recognition results of the five combined actions are shown in Table 2.

Compared with Table 1, the recognition accuracy rate of the five kinds of actions in Table 2 has greatly been improved, indicating that the method for identifying human action is very effective. Wherein the recognition rate for the static actions, say, the right arm lifting, squatting, is relatively high because the characteristics of static action is fixed, action characteristic curve is small. The training feature set can recognize this action very well and action recognition efforts can well exclude similar static action interference. As for dynamic movements, say, the recognition rate for upper limb external rotation is relatively low because when the same person or different people do the same action, the rate will be different and the action characteristic curve length will be different. The action characteristic curve length adopted in training and recognition in the present paper is one with a fixed length, and the recognition rate is vulnerable to stretching or compressing of the action characteristic curve length.

Our discussion on the sample training and the identifications of body movements is based on the Adaboost algorithm. Compared with other methodologies, the present methodology puts forward temporal and spatial classification features and uses SVM algorithm to get weak classifiers selecting human action set depending on different human action features, which is simple to achieve and of high operating efficiency. Experimental results show that this method has a high recognition rate. Human skeleton data herein is obtained by Kinect deep sensor. Kinect deep sensor can overcome the influence of light and other environmental factors with relatively low pricing and suitability for human action identification. The use of bone rotation information for action recognition, showing a valid point and the position of the bone change over the bone can provide more data for the physical rehabilitation and action recog-

Table 2. Recognition rate of five combined actions

Action	Output the right arm lifting	Output upper limb external rotation	Output the right leg flat lifting	Output the right leg side lifting	Output squatting	Recognition failure	Accuracy rate
the right arm lifting	192	3				4	96%
upper limb external rotation	4	188				8	94%
the right leg flat lifting			191	4		5	95.5%
the right leg side lifting			3	193		4	96.5%
squatting					194	6	97%

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

nition which can effectively avoid the normalization and gimbal lock problems and can improve the recognition rate of movements, especially the distinction and recognition between two similar actions. Many difficulties are also encountered during the identification process, such as the selection of action characteristic curve length, and the calculation of curve similarity. Here, we successfully resolved all the above-mentioned problems through demonstrated research and experiments.

4. KINECT-BASED AND 3D VIRTUAL SIMULATION LIMB REHABILITATION TRAINING

Many defects exist in traditional rehabilitation. The introduction of virtual reality (VR), which in addition to real experience and way of thinking, brings the patients into the virtual realm as a complete individual. In such an environment, the patient's visual, auditory and tactile senses can all be fully capitalized. Compared with conventional treatments, VR technology can provide a more realistic, more systematic and natural environment, aside from improved scientific training intensity and more meaningful assessments for patients.

There is many software of 3D environment visualization technology. With the birth of Kinect somatosensory deep cameras, Kinect 3D plug-in software subsequently emerged. Kinect in health informatics context can capture real-time characteristics of the human skeleton, the characters and 3D animated characters are bound to achieve high level human-computer interactions (HCIs). Kinect-Unity-interactive KinectForWindows2.0_Unity.unitypackage plug–in software, the official plug-in of Unity, which can be simply deployed for limb rehabilitation in 3D scenes.

System design is divided into Unity3D scene display module, Unity3D interface module and Kinect data acquisition module. The scene module function can realize the character model to change with the human body captured by the lens. Model binding effect is shown in Figure 5.

As can be seen, the software application allows the setting of a character in the scene, making a character in 3DMax, as well as attaching a character bone, binding in Unity3D, and adjusting the light and sky. Using Unity3D interface modules, the post-Tim script, corresponding with the properties of the front Unity3D, one can achieve bone binding, mirror movement, close-range mode, smoothing and so forth. Additionally, the scene module function can realize the character model to change with the human body captured by the lens. Body parts' binding facilitates the skeleton data acquisition in the Unity3D scene. Many groups of bones are used to represent the various bone nodes in the scene to complete the mapping function of the bones. In transforming the skeletal visualization, we use the four-element approach.

The four-element approach is simply hyper-complex. Four elements are composed of real numbers and linear combinations of *i, j, k*, yielding the formula as follows:

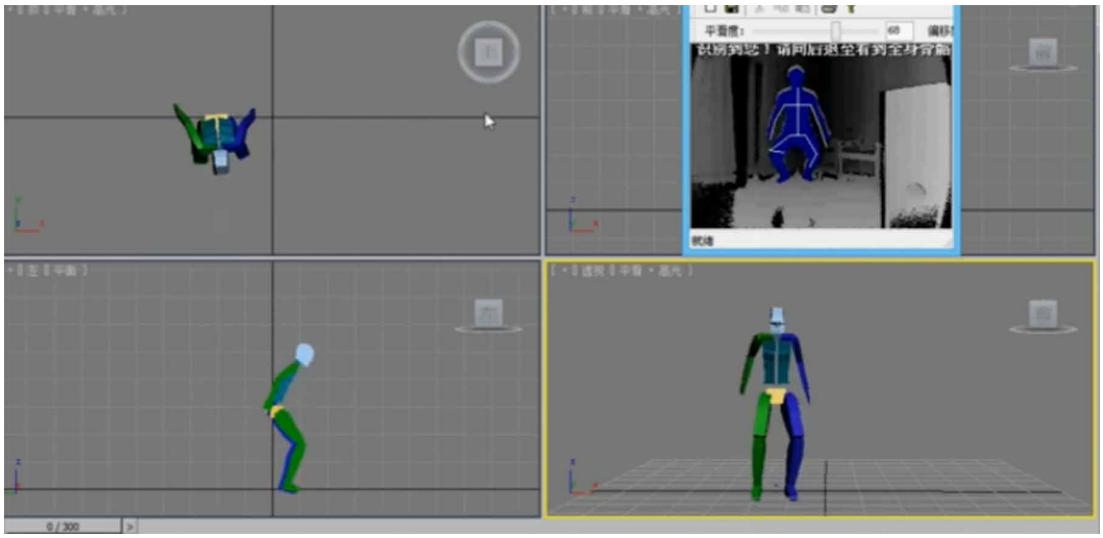
$$i_2 = j_2 = k_2 = -1 \tag{16}$$

$$a + b_i + c_j + d_k = 0 \tag{17}$$

In the computation of the rotation angle, it is worth noting that the four-element Quaternion is used. Quaternion variables are relatively small. Let *x, y, z, w* take values on [-1,1]. The object rotates for two weeks and all its values return to the initial value. The initial value: (0,0,0,1), if it rotates along the y-axis of 180 °, 360 °, 540 °, 720 °, the value taken corresponding to the angle is (0,1,0,0),(0,0,0,-1),(0,-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Human model binding effect



1,0,0),(0,0,0,1); if it rotates along the x-axis of 180°, 360°, 540°, 720°, the value taken corresponding to the angle is (-1,0,0,0), (0,0,0,-1), (1,0,0,0), (0,0,0,1).

No rotation is written by Quaternion. Identify in this way, the bone binding, self-mapping and bone rotation can be achieved.

Kinect data are divided into three parts: device control, bone binding, device image acquisition. The control of the device is mainly reflected in: device starting and initialization, camera angle adjustment, device turning off. Smoothing facilitates free bone rotation, prevents skipped and teleported effects, and is implemented in the bone rotation algorithm.

The final procedure is the test of each module. Limb recognition is the identification of human limbs by mathematical algorithm. Movement identification can be from various parts of the human body, but usually refers to the movement of the head and limbs. You can use Kinect to collect the depth image to do segmentation, confine the range of depth and extract the image within the range. This way of achieving body recognition, as well as recognition speed and accuracy has been greatly improved versus that of the conventional identification methods.

As depicted in Figure 6, here are two games for limb rehabilitation training based on Kinect somatosensory technology and 3D scene technology.

After the scientific research of Kinect-based rehabilitation system, we focus on a comparative survey of the research results and the ways of traditional limb rehabilitation of patients. The investigation results are shown in Table 3:

From the survey results, it is apparent that the limb rehabilitation training method based on Kinect and 3D virtual simulation technology is significantly more efficient than the traditional methods of limb rehabilitation in time use efficiency, training improvements and fun.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Games for limb rehabilitation based on Kinect somatosensory & 3D scene technology

Table 3. Research results of limb rehabilitation training based on Kinect & 3D VR technology

Content	Research group(n = 20)		Control group(n = 20)	
	satisfaction	dissatisfaction	satisfaction	dissatisfaction
valid training time	17(85)	3(15)	3(15)	17(85)
Improvement of purposes and interests of limb training	19(95)	1(5)	11(55)	9(45)
Improvement of the initiative and enthusiasm of limb training	18(90)	2(10)	5(25)	15(75)
Improvement of clinical competence	16(80)	4(20)	5(25)	15(75)
improvement of doctor-patient communicative competence	17(85)	3(15)	8(40)	12(60)
preference for this limb rehabilitation	17(85)	3(15)	9(45)	11(55)

5. CONCLUSION

Traditional rehabilitation needs to be carried out in a formal clinical setting such as hospitals. Doctors are basically needed to train patients face-to-face with serious movement disorders.

In contrast, this work proposes a series of upper limb rehabilitation training methods and systems based on Kinect within the emerging health informatics context. It first gathers standardized training actions of doctors, establishes the standard rehabilitation action database, and then identifies and matches exercise postures through joint points and bone data with the help of Kinect equipment. In this context, patients are empowered and can be asked to do the appropriate rehabilitation exercises according to different symptoms, which can be done at home only with the help of inexpensive computer, and other relevant display devices. Such innovations will greatly reduce the workload of doctors.

As well, the exercise interests of the patients trained in the 3D virtual scene can be greatly improved. Doctors can assess the effect of rehabilitation by matching the exercise data of the rehabilitation coach and rehabilitation patients to improve the efficiency and accuracy of the identification operation by AdaBoost algorithm. Experiments show that the model principle is practical, with the platform having commercial value.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In summary, many attempts of Kinect exercise rehabilitation, due to the limitations of Kinect equipment, have been quite limited to date. Oftentimes, there exists error range in skeleton measuring to make the existing skeleton tracking technology incapable in term of expressing human actions, especially human complex continuous actions. Accordingly, in light of advances in health informatics, it is now possible to improve the tracking accuracy of Kinect human skeleton and the algorithm of motion recognition, so as to provide better and more natural human-computer interactions.

ACKNOWLEDGMENT

This research was supported by Project of the Education Department of Jilin Province (Project No. 2015-LY-5-46-L48) and Jilin Science and Technology Development Plan Project (Project No. 20170204033GX).

REFERENCES

Amami, R. (2013). *The challenges of SVM optimization using Adaboost on a phoneme recognition problem. In Cognitive Infocommunications* (pp. 25–28). Seoul: IEEE.

Chao, Z., Zhao, X., & Su, P. (2016). Research on EEG Recognition Algorithm Based on SVM Classifier. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, *9*(12), 199-206. doi:10.14257/ijsip.2016.9.12.19

Ji, Z., & Bai, D. (2016). Research of Optimal path Automatic Navigation in the Prescribed Scope. *International Journal of Multimedia and Ubiquitous Engineering*, *11*(12), 229-236. doi:10.14257/ijmue.2016.11.12.21

Lai, K., Konrad, J., & Ishwar, P. (2012). A gesture-driven computer interface using Kinect. In *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, Shenzhen (pp. 11-14).

Lin, T. Y., Hsieh, C. H., & Lee, J. D. (2013). A Kinect-based system for physical rehabilitation: Utilizing tai chi exercises to improve movement disorders in patients with balance ability. In *Proceedings of the 7th Asia Modelling Symposium (*pp. 10-13). IEEE Computer Society. 10.1109/AMS.2013.29

Maekawa, T., Yanagisawa, Y., Kishino, Y., Ishiguro, K., Kamei, K., Sakurai, Y., & Okadome, T. (2010, May). Object-based activity recognition with heterogeneous sensors on wrist. In Pervasive Computing (pp. 246–264).

Raptis, M., Kirovski, D., & Hoppe, H. (2011). Real-time classification of dance gestrures from skeleton animation. In *Proceedings of the ACM SIGGRAPH, Eurographics Symposium on Computer Animation* (pp. 16-18).

Xie, L., Liao, H., & Yang, Y. (2013). Recognition and application research of Kinect-based gesture (pp. 258–260). Computer Technology and Development.

Shotton, J., Fitzgibbon, A., & Cook, M. (2011). Realtime Human pose recognition in parts from single depth images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, New York (pp. 15-19).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yang, N.-E., Kim, Y.-G., & Park, R.-H. (2012). Depth Hole Filling Using the Depth Distribution of Neighboring Regions of Depth Holes in the Kinect Sensor. In *Proc. of the IEEE Conf. Signal Processing, Communication and Computing (ICSPCC)* (pp. 658-661). 10.1109/ICSPCC.2012.6335696

Ye, Z., Li, Y., & Han, B. (2014). Appearance Design Evaluation on CNC Machine Tools Based on Fuzzy Synthetic Evaluation Model. *International Journal of Multimedia and Ubiquitous Engineering*, *9*(2), 65–76. doi:10.14257/ijmue.2014.9.2.07

Yim, H.-B., & Lea, S.-R. (2014). Compensation Characteristics of the Distorted WDM Signals Depending on the Variation Degree of RDPS Random Variables in the Optical Links with Dispersion Management and Optical Phase Conjugation. *International Journal of Control and Automation*, *7*(5). doi:10.14257/ijca.2014.7.5.01

Yuzhen, C., Weichao, C., & Yang, C. (2010). *Body posture detection technique based on MEMS acceleration sensor*. Nanotechnology and Precision Engineering, 8(1), 37–41.

Zhou, M., Zhao, L., Zhang, Y., Gao, Z., & Pei, R. (2015). Pure Electric Vehicle Power-train Parameters Matching based on Vehicle Performance. *International Journal of Control and Automation*, *8*(9), 53-62. doi:10.14257/ijca.2015.8.9.06

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 13(3); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 49-64; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 53
Knowledge Acquisition Through Ontologies from Medical Natural Language Texts

José Medina-Moreira
University of Guayaquil, Ecuador

Oscar Apolinario-Arzube
University of Guayaquil, Ecuador

Katty Lagos-Ortiz
University of Guayaquil, Ecuador

María del Pilar Salas-Zárate
University of Murcia, Spain

Harry Luna-Aveiga
University of Guayaquil, Ecuador

Rafael Valencia-García
University of Murcia, Spain

ABSTRACT

Ontologies are used to represent knowledge and they have become very important in the Semantic Web era. Ontologies evolve continuously during their life cycle to adapt to new requirements and needs, especially in the biomedical field, where the number of ontologies and their complexity have increased during the last years. On the other hand, a vast amount of clinical knowledge resides in natural language texts. For these reasons, building and maintaining biomedical ontologies from natural language texts is a relevant and challenging issue. In order to provide a general solution and to minimize the experts' participation during the ontology enriching process, a methodology for extracting terms and relations from natural language texts is proposed in this work. This framework is based on linguistic and statistical methods and semantic role labeling technologies, having been validated in the domain of diabetes, where they have obtained encouraging results with an F-measure of 82.1% and 79.9% for concepts and relations, respectively.

DOI: 10.4018/978-1-7998-1204-3.ch053

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

In the biomedical context, where the generation of new knowledge is continuous, ontologies have raised interest in the entire biomedical community (Rubin, Shah, & Noy, 2008). Ontologies are used for organizing and sharing biological knowledge as well as integrating different sources of knowledge in order to provide interoperability among different research communities. For example, MEDLINE system, which is based on the MESH medical ontology, is the largest medical knowledge base available over the Internet. Due to the fact that in medical domain new terms are continually emerging, the update and extension of MESH is necessary to avoid its obsolescence. In the last years, ontologies have been applied to several areas such as natural language processing (Paredes-Valverde, Valencia-García, Rodríguez-García, Colomo-Palacios, & Alor-Hernández, 2015), sentiment analysis (Salas-Zárate, Valencia-García, Ruiz-Martínez, & Colomo-Palacios, 2016), bioinformatics (Mayor & Robinson, 2014), business intelligence (Rodríguez-García et al., 2016), recommender systems (Colombo-Mendoza, Valencia-García, Rodríguez-González, Alor-Hernández, & Samper-Zapater, 2015), and cloud computing (Rodríguez-García, Valencia-García, García-Sánchez, & Samper-Zapater, 2014a) (Rodríguez-García, Valencia-García, García-Sánchez, & Samper-Zapater, 2014b), to mention a few.

Nevertheless, building and maintaining long-living ontologies is still a challenge for the scientific community, which requires a considerable human effort and development time. An ontology lifecycle includes several phases such as creation, maintenance, evaluation, consistency checking, etc. and there are several ongoing attempts to address these issues by providing automatic ontology management methods (Petrucci, Ghidini, & Rospocher, 2016). Among them, several ontology learning and ontology evolution methods have been developed.

Ontology learning has been an active research area that has grown according to the needs of creating ontologies in a fast and reliable way. Ontology Learning (also named ontology generation or ontology extraction) is a knowledge acquisition activity that relies on (semi-) automatic methods to transform unstructured (e.g. corpora), semi-structured (e.g. folksonomies, html pages, etc.) and structured data sources (e.g. databases) into conceptual structures (Petrucci et al., 2016).

On the other hand, ontology evolution (also named ontology enrichment) can be seen as the activity of enriching ontologies by adding or modifying concepts and relations. More specifically, in (Stojanovic, Maedche, Motik, & Stojanovic, 2002) ontology evolution is defined as the timely adaptation of an ontology to modified business requirements, to trends in ontology instances and patterns of usage of the ontology-based application, as well as the consistent management/propagation of these change to dependent elements.

Ontology evolution is considered a process due to the fact that a change in the ontology can cause inconsistencies in other parts of the ontology, as well as in other dependent applications (Abcckcr & Stojanovic, 2005), for that reason, the process of ontology evolution contemplates some tasks such as for example resolving the ontology changes, ensuring the consistency of the underlying ontology and facilitating the changes and management of the ontology to the user (Stojanovic et al., 2002).

This paper focuses on the tasks of extracting terms and semantic relationships from natural language texts in order to simplify the process of building and updating the concepts and relationships in biomedical ontologies. The proposed framework is based on linguistic and statistical term extraction methods and semantic role labeling technologies, having been validated in the diabetes domain. Many biomedical ontologies are structurally quite simple, since they consist of only a few relationship types apart from the taxonomical relationship that constitute the backbone of an ontology. One of the main contributions of this paper is the provision of a method for increasing the typology of semantic relationships.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The structure of this paper is described as it follows. In section 2, a summary of the most important ontology learning approaches is explained. Then, in Section 3 the ontology evolution approach is described in detail. Next, in Section 4 an experiment and its results are shown. Finally, some conclusions and future work are put forward in the last section.

RELATED WORK

In the last years, several approaches and tools have been proposed to address the ontology learning and ontology evolution topics. The main advances in the ontology learning are to find and identify ontology elements such as concepts, instances, relations. These elements are mainly identified semi or automatically through approaches based on Natural Language Processing (Sinha, Garg, & Chandra, 2016) (de Azevedo et al., 2014), pattern recognition (Chen, Chen, Leong, & Jiang, 2013), and data mining (Guo, Schwartz, Burstein, & Linger, 2015), among others.

Significant approaches for Ontology learning have arisen from Ontology evolution, which are focused mainly on the management of the evolution process, and versioning comparison. For example, in (Maedche, Motik, Stojanovic, Studer, & Volz, 2002) one of the first methodologies for integral ontology evolution is proposed. Their approach is based on six-phase evolution process, which analyses the causes and the consequences of the changes and ensures the consistency of the ontology and depending artifacts after resolving these changes. This approach was implemented in the KAON ontology management system. Other approach is (Klein, Fensel, Kiryakov, & Ognyanoff, 2002) which affords the topic of ontology versioning for web-based ontologies. The system is based on the comparison of RDF-based ontologies in three different dimensions, meta-data, conceptual relations and transformations between the ontologies themselves. The integration of different ontologies is also afforded in OnEX (Hartung, Kirsten, Gross, & Rahm, 2009). It uses existing ontology versions and identifies the differences between them in order to integrate different versions of the same ontology. The evolution mapping consists of adding, deleting and updating, concepts, relations, and attributes from the last version of the ontology. In (Gil & Martin-Bautista, 2014) they introduce a Systematic Methodology for Ontology Learning (SMOL) from heterogeneous sources through seven phases (Methodology strategy selection, knowledge discovery, query requirements, knowledge selection, knowledge structure construction, knowledge exploration and search, and knowledge structure reorganization) over structured flow. In addition, modification and versioning have been considered for the workflow design of SMOL. Finally, aiming to validate the SMOL methodology an academic case study was carried out. Meanwhile, in (Dinh, Reis, Pruski, Silveira, & Reynaud-Delaître, 2014) a method to recognize the change of patterns between ontology versions is proposed. In order to evaluate the effectiveness of their proposed method, the authors carried out experiments by using several versions of biomedical ontologies, namely SNOMED-CT (SCT), MeSH and ICD-9-CM (ICD9).

However, few studies deal with the automatic evolution of ontologies from texts. For example, in (Zablith, Sabou, d'Aquin, & Motta, 2009) presented the Evolva framework. The authors propose an approach that tries to solve two functionalities, the adaptation of ontologies and change management. Meanwhile, the study presented in (Ottens, Hernandez, Gleizes, & Aussenac-Gilles, 2009), proposed the Dynamo (Dynamic ontologies) tool which used an adaptive multi-agent system to represent the ontology itself and to produce the ontology. In order to evaluate the quality of a taxonomy built with Dynamo, the authors considered a reference taxonomy constructed from a corpus named Astro. The

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

corpus consists of a collection of more than 3000 summaries of scientific articles published in 1995 and 2002 in the journal "Astronomy & Astrophysics". More recently the authors have proposed an improved version named Dynamo MAS (Sellami, Camps, & Aussenac-Gilles, 2013). This tool is based on an adaptive multi-agent system (adaptive MAS or AMAS) that allows building and to evolve ontologies from the text. Lexico-syntactic patterns were used to extract lexical relations between terms. Also, the authors carried out several experiments to evaluate the quality of the system on three corpora, namely Artal (287 bug report documents), Arkeotek (299 rule-based formulation of scientific papers), and Actia (documents that report fault descriptions and repair procedures). Final results showed that the quality of Dynamo-MAS depends on the corpus and initial state of the ontology. In (Benomrane, Sellami, & Ben Ayed, 2016), OntoAMAS, an ontologist feedback tool based on an adaptive multi-agent system (adaptative MAS or AMAS) is presented. OntoAMAS consists of four modules: 1) addition of new documents, 2) extraction and filtering of terms, 3) knowledge interpretation, ontology update, and 4) ontologist interventions. Aiming to evaluate the feasibility of their approach, a set of experiments were carried out, obtaining encouraging results.

Unlike these systems, our proposal is not based on the integration of different versions of the same ontology, but in the acquisition of new concepts and relationships from textual sources. Any seed ontology from the domain under question can be updated and expanded using as input the textual resources available into the scientific literature repositories.

ONTOLOGY EVOLUTION PROCESS

In this work, an ontology is seen as a formal and explicit specification of a shared conceptualisation (Studer, Benjamins, & Fensel, 1998). The Web Ontology Language (OWL) has been used to model the biomedical ontologies to be evolved from natural language texts. In OWL, the following main ontological entities can be distinguished: classes, subclass of relationships, datatype properties (attributes), object properties (relationships), and individuals. OWL provides a formal theory for taxonomy, whereas any other semantic relationship, like partonomy or topology, must be manually defined and implemented using object properties for that.

Our ontology evolution approach deals with the addition and removal of concepts and relationships from a domain ontology. These ontology entities (concepts and relationships) are obtained from natural language texts through two main subprocesses (i) the term extraction process and (ii) the relation extraction process. Once the relevant elements are obtained, the ontology is expanded or modified by using (iii) the ontology evolution module. The main premise of any ontology evolution approach is that the ontology remains consistent after applying a change. In this approach, the consistency of the ontology is checked after each change using and OWL2 reasoner.

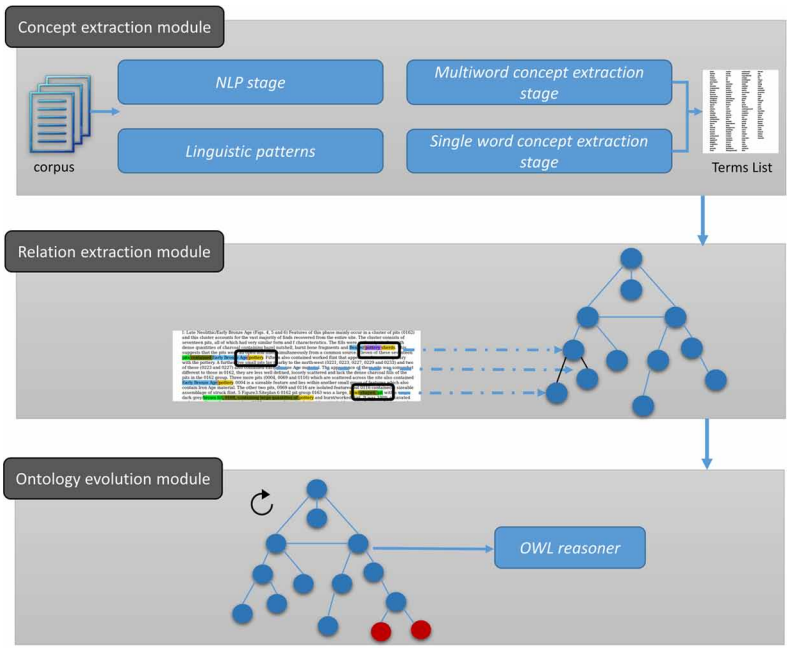
Next, in Figure 1 the main modules of the architecture are described in detail.

Concept Extraction Module

First of all, terms representing concepts are identified. Automatic term recognition denotes the process used to systematically extract pertinent terms and their variants from a collection of documents (Ananiadou & McNaught, 2006). It is assumed that there exist both multiword and single word terms. By taking into account this assumption, two different methods have been implemented: the NC-Value

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. System architecture



algorithm, which allows obtaining the multiword terms candidates to represent concepts, and TF-IDF which has been employed to obtain terms formed by one word. This process can be decomposed into several phases as it is described below.

NLP Stage

The main aim of this stage is the extraction of the morphosyntactic structure of each sentence. For this purpose, the GENIA tagger (Tsuruoka et al., 2005) integrated into GATE has been used. The GENIA tagger analyzes English sentences and outputs the base forms, part-of-speech tags, chunk tags, and named entity tags. The tagger is specifically tuned for biomedical text such as MEDLINE abstracts.

On the other hand, GATE (Cunningham, 2014) is an infrastructure for developing and deploying software components that process human language.

Linguistic Patterns

The candidate terms are identified by means of a hybrid method which uses a series of linguistic patterns in which the morphosyntactic structure of the terms is described. The use of linguistic patterns for identifying biomedical terms is a difficult issue due to the particularities of the biomedical domain, for example, the use of non-alphanumeric characters and acronyms, the complexity of some compound names, such as “granulocyte-macrophage colony-stimulating factor” (Ananiadou & McNaught, 2006). Indeed, terms representing biomedical concepts are usually comprised of multiple words that lead to a more specific meaning as a whole. For instance, “virus” has a more general meaning than “simian immunodeficiency virus”.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. An example of some linguistic patterns

Term	Linguistic Pattern
human lymphocyte line	[JJ] [NN] [NN]
synthesis of 2Fe-2S cluster	[NN] [IN] [NN]
peri-kappa B-site	[NN] [NN]
immunoregulatory protein	[JJ] [NN]

In order to address this issue, relevant biomedical patterns have been obtained from some PubMed (United States National Library of Medicine, 1996) abstracts, which were analyzed using GENIA tagger. The most frequent patterns were collected in order to discover new terms.

Table 1 shows some of the morphosyntactic patterns obtained for the biomedical domain as well as some terms matching the patterns.

Multiword Concept Extraction Stage

Once a list of multiword candidate terms has been obtained from the linguistic patterns, it is filtered out by applying the NC-value algorithm presented in (Ochoa, Martínez, Almela, & Valencia-García, 2010). In this algorithm, the system processes context words and split them up according to their grammatical category (i.e., Adjectives, Verbs or Nouns). Both the C-Value and NC-Value algorithms are explained in detail.

C-Value

First, the system arranges the term list and calculates the frequency of occurrence of the candidate term within longer candidates, the frequency of occurrence of that longer candidate term, the length of the candidate term and the total frequency of occurrence of the candidate term in the corpus (see equation 1).

$$C-value = \begin{cases} log_2 \left|a\right| * f\left(a\right) \\ log_2 \left|a\right| * \left(f\left(a\right) - \frac{1}{P(T_a)} \sum_{b \in T_a} f\left(b\right) \right) \end{cases} \quad (1)$$

where:

a is the candidate string;
|a| is the length of the candidate string;
f(a) is its frequency of occurrence in the corpus;
Ta is the set of extracted candidate terms containing a;
P(Ta) is the number of the longest candidate terms containing a;
Sf(b) is the frequency of occurrence of a sub-term of any candidate term b as |a| < |b|.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

NC-Value

In order to obtain an acceptable degree of accuracy in the candidate term list, the NC-value method uses the morphological information from the context of the term under question. The system processes context words and splits them up according to their grammatical category of Adjectives, Verbs or Nouns. With the method developed by (Grefenstette, 2012), a type of weight known as 'context weighting factor' is obtained. It calculates the probability of a context word appearing with a term, as it is formalized in equation 2:

$$weight\left(w\right) = \frac{t\left(w\right)}{n} \tag{2}$$

where:

w is the context word;
t(w) is the number of times that the context word appears with the term;
n is the total frequency of occurrence;
weight (w) is the context weighting factor.

With the weight score obtained, the NC-value may by calculated by using the equation 3:

$$NC - value\left(a\right) = 0.8C - value\left(a\right) + 0.2\sum_{b \in C_a} f_a\left(b\right) weight\left(b\right) \tag{3}$$

where:

a is the candidate term;
Ca is the set of context words for a;
b is a word from Ca;
fa (b) is the frequency of b as the context word for a;
weight (b) is the weight of b as the context word.

In equation 4, the factors 0.8 in the first part and 0.2 in the second part have been assigned after several tests conducted by (Grefenstette, 2012). If the C-value of the candidate term is multiplied by 0.8 and the result of the summation of the individual weight corresponding to the context words is multiplied by 0.2, the NC-value is obtained, which rearranges the list by placing the best candidate terms on the top of the list.

Single Word Concept Extraction Stage

The single word concept extraction is based on the TF-IDF algorithm (Salton and McGill, 1983), which uses the following equation (see equation 4).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\left(tf-idf\right)_{i,d}=\frac{n_{i,d}}{\sum_k n_{k,d}}*\log\frac{\left|D\right|}{N_i} \tag{4}$$

where $n_{i,d}$ is the number of occurrences of the ontological entity i in the document d, $\sum_k n_{k,d}$ is the sum of the occurrences of all the ontological entities identified in the document d, $|D|$ is the set of all documents and N_i is the number of all documents annotated with i.

Relation Extraction Module

In the previous phase, an ordered list with the terms that represent the concepts of the domain has been obtained. For each term, a concept is identified and in this module, the semantic relations between these concepts have to be obtained. Relationships between entities in natural language texts are primarily expressed as verbal relationships or by means of their corresponding nominal form (Friedman, Kra, & Rzhetsky, 2002). The semantic relationship between a syntactic constituent, for example, a noun, and a predicate, for example, verb, is called semantic role. A semantic role defines the role of a verbal argument in the event expressed by the verb (Moreda, Llorens, Saquete, & Palomar, 2011).

For example, in the sentence 'Gene MCOLN1 is mutated in Mucolipidosis type IV', the verb expression "is mutated" express the relation between the events or arguments, which are defined as "Gene MCOLN1" and "Mucolipidosis type IV". Semantic roles and concept labels or tags are also assigned to these arguments. For example, the verb "mutate" expects to have at least two arguments, (i) an initial entity, in this case, "Gene MCOLN1 and" (ii) the final entity, that is, the changed entity, in this case, "Mucolipidosis type IV".

In the biological domain, semantic roles have been used for tasks such as extracting semantic relations and Named Entities (NEs). More concretely, there are some domain-specific projects where semantic roles are defined for biomedical verbs, such as PASbio (Wattarujeekrit, Shah, & Collier, 2004), BioFrameNet (Dolbey, Ellsworth, & Scheffczyk, 2006) and BIOSMILE (Tsai et al., 2007).

In this approach, a framework based on previous works (Ruiz-Martínez, Valencia-García, Martínez-Béjar, & Hoffmann, 2012) is used. This framework attempts to support domain experts in populating ontologies from natural language texts. It allows for several semantic relations to be used and reduces the degree of expert participation during the ontology construction process. It is based on the association of semantic frames from FrameNet with an ontological model created from top-level ontologies. In the BioOntoVerb framework, there are 10 top-level relationships which represented a wide range of biomedical predicates. These predicates relate two ontological concepts on a textual level, although they are constrained to the top-level relationships. For this reason, new predicates have been added, even if they have not a correspondence with top-level relationships.

In Table 2 an excerpt of the relationships described in BioOntoVerb framework is presented.

This module works as follows. First, the concepts obtained in the previous phase are identified in each sentence in the corpus. Then, the verbal expression of the current sentence is found and the system searches if there is any semantic frame that could be applied in order to extract a relationship between the concepts identified in this sentence. If so, a semantic relationship between the two concepts is extracted and included in the domain ontology.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. An excerpt of the mapped relationships

Ontological Relationship	FrameNet Frames	Lexical Units	Examples in Corpora
caused_by	Cause: A Cause causes an Effect	Cause, Induce, Make, Create, Incite, Provoke, Engender, Instigate, Generate, Produce, Result, Develop	nur77 *may cause* apoptosis by monomer K30p (p13K30) and K34p (p13K34) *result in* a Trp - Arg substitution
located_in	Locating: A Perceiver determines the Location of a Sought_entit	Locate, Situate, Place, Localize, Localized in, Be present In, Localized to, Find	One type, EA / S, *is located in* the upstream promoter region I kappa B gamma - 1 *is found in* both the cytoplasm and nucleus
contained_in	Contain: Container holds within its physical boundaries the Contents	Create, Originates from, Derive from, Generate (inversa), Come from, Develop from, Alter, Transcribe from, Stems from	cytotoxicity of indolocarbazoles *derives from* Top1p p45 mRNA *is transcribed only from* aNF - E2 promoter
transformation_of	Transform: An Agent or Cause causes an Entity to change.	Transform, Modify, Change Transform in/into, Changes in/ into Mutate in/into, Changes into, Converts into, Turns into	I kappa B alpha *is* selectively *modified in* LMP - 1 - expressing B cells GATA sites *have been mutated into* human T lymphocytes

Ontology Evolution Module

The main objective of this module is to manage the evolution of the domain ontology. The main changes in the ontology are mainly the addition and removal of ontological entities, mainly concepts, and relationships.

For this, when a new document is added to the corpus, this module is in charge of launching both the concept and relation extraction modules. If new ontological entities are identified they are inserted in the ontology. On the other hand, if a document is removed the module only will launch again the concept extraction module in order to check if any concept should be removed from the ontology. The relationships are related to the documents from they were extracted, if the document that contained a relationship is removed, then the relationship should be also removed from the domain ontology.

During the evolution process, an OWL reasoner is used to detect inconsistencies in the ontology. If any inconsistency is identified, then the system alerts to the user that an inconsistency has been produced due to an add/remove ontological entity action.

VALIDATION

The methodology described in this work has been validated across the diabetes domain. The study of the results is based on three main scores: precision, recall, and F1-Measure. These measures are the most commonly used for the assessment of statistical extraction systems, and trace their origins back to IR discipline. The precision score is the result of the division between the amount of ontological entities suggested by the system and that are accepted by the expert user, and the total amount of ontological entities suggested. The recall score is the result of the division between the amount of ontological entities suggested by the system and accepted by the expert, and the total amount of ontological entities existing

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in the fragment. Finally, the F-measure score can be interpreted as a weighted average of the precision and recall, where an F-measure score reaches its best value at 1 and worst score at 0.

The experimental corpus in the diabetes domain has 174,568 words distributed into 125 documents. This corpus has been manually processed by domain experts, so obtaining a total amount of 834 concepts or classes and 987 semantic relationships.

As can be observed in Table 3, the system identified 895 concepts and 987 relations, of which the expert accepted a total of 710 concepts and 760 relations.

The results seem promising, obtaining a precision score of 79.3%, a recall score of 85.1% and an F-measure score of 82.1% for concepts detection, and a precision score of 82.6%, a recall score of 77.0% and an F-measure score of 79.7% for relations detection. Therefore, the results show that the system is able to propose several concepts and relations that the experts have manually detected.

Furthermore, a scenario has been defined in order to evaluate the ontology evolution process. In this case of study, 30 new documents were included in the corpus. This corpus was analyzed by the expert user to detect domain concepts and relations. A total of 35 new concepts and 16 new relations were obtained.

The same corpus was processed by the system obtaining 30 new concepts and 12 new relations, of which 26 concepts, and 10 relations were accepted by the expert user. Results are showed in detail in Table 4.

The system obtained good results for ontology evolution process with a precision of 68.4% and 71.4% for concepts and relations, respectively. However, the authors consider that the detection of terms can be improved through the inclusion of other more complex linguistic patterns that allow the detection of some concepts that have not been identified.

Table 3. Evaluation results in the diabetes domain

	Concepts	**Relations**
Relevant concepts and relations	834	987
Concepts and relations identified by the system	895	920
Concepts and relations accepted by the expert	710	760
Precision	0.793	0.826
Recall	0.851	0.770
F-measure	0.821	0.797

Table 4. Results after adding 30 new documents to the corpus

	Concepts	**Relations**
Relevant concepts and relations	35	16
Concepts and relations identified by the system	38	14
Concepts and relations accepted by the expert	26	10
Precision	0.684	0.714
Recall	0.742	0.625
F-measure	0.712	0.666

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

In this paper, an automatic method for acquiring knowledge from biomedical texts has been presented. This approach is based on the use of semantic roles in order to extract semantic relations between concepts.

The proposal supports an incremental evolution by taking into account new text documents. Ontology evolution from free text documents is an important activity for the knowledge engineering community. The approach presented in this work is totally automatic. Another key feature of this approach is that it works not only with taxonomies, but also with multiple semantic relations.

Similarly, (Sánchez & Moreno, 2008) presents a methodology for the detection of non-taxonomic relations from texts. It is based on the identification of the relevant verbs in the text chunks. These verbs are then used as a knowledge basis for learning and tagging non-taxonomic relations automatically and without supervision. Both studies use linguistic patterns for obtaining taxonomic relations.

The authors in (Jiang & Tan, 2010) introduce a methodology which has a similar common ground to ours. The aim of these authors is the development of a high-quality ontology, and for this purpose they use a combination of statistical and lexical-semantic methods.

Further validations of the system are planned by means of its application to texts from different medical domains and by using statistical methods for the analysis of the results obtained. Moreover, the authors intend to extend the system to cover special and temporal relationships (Cuadra, Calle, & Rivero, 2011) and axioms (Terrientes, Moreno, & Sánchez, 2010). The main forecast problem concerning axioms is, however, that the number of participants is a priori unknown. Notwithstanding this fact, the number of axioms present in a text is irrelevant in comparison to the number of other knowledge entities.

It is also planned to include some methods to measure the quality of the acquired ontologies such us the ones presented in (Duque-Ramos, Fernandez-Breis, Stevens, & Aussenac-Gilles, 2011) (Duque-Ramos et al., 2013) and to introduce some methods of question answering in the evolutioned ontologies such as the one presented in (Vila, Mazon, & Ferrandez, 2011).

ACKNOWLEDGMENT

This work has been funded by the Universidad de Guayaquil (Ecuador) through the project entitled "Tecnologías inteligentes para la autogestión de la salud". María del Pilar Salas-Zárate is supported by the National Council of Science and Technology (CONACYT), the Secretariat of Public Education (SEP) and the Mexican government. Finally, this work has been also partially supported by the Spanish Ministry of Economy and Competitiveness and the European Commission (FEDER / ERDF) through project KBS4FIA (TIN2016-76323-R).

REFERENCES

Abcckcr, A., & Stojanovic, L. (2005). Ontology Evolution: MEDLINE Case Study. In P. D. O. K. Ferstl, P. D. E. J. Sinz, D.-W. I. S. Eckert, & D.-W. I. T. Isselhorst (Eds.), *Wirtschaftsinformatik 2005* (pp. 1291–1308). Physica-Verlag HD. doi:10.1007/3-7908-1624-8_68

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ananiadou, S., & McNaught, J. (2006). *Text Mining for Biology and Biomedicine*. London: Artech House Publishers.

Benomrane, S., Sellami, Z., & Ben Ayed, M. (2016). An ontologist feedback driven ontology evolution with an adaptive multi-agent system. *Advanced Engineering Informatics*, *30*(3), 337–353. doi:10.1016/j.aei.2016.05.002

Chen, X., Chen, C.-H., Leong, K. F., & Jiang, X. (2013). An ontology learning system for customer needs representation in product development. *International Journal of Advanced Manufacturing Technology*, *67*(1–4), 441–453. doi:10.100700170-012-4496-2

Colombo-Mendoza, L. O., Valencia-García, R., Rodríguez-González, A., Alor-Hernández, G., & Samper-Zapater, J. J. (2015). RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes. *Expert Systems with Applications*, *42*(3), 1202–1222. doi:10.1016/j.eswa.2014.09.016

Cuadra, D., Calle, J., & Rivero, J. (2011). How to manage spatio-temporal events in relational databases. *Journal of Research and Practice in Information Technology*, *43*(4), 329.

Cunningham, et al. (2014). *Developing Language Processing Components with GATE Version 8*.

de Azevedo, R. R., Freitas, F., Rocha, R. G. C., de Menezes, J. A. A., Rodrigues, C. M. O., & Gomes, M. C. (2014). Towards a Framework for Ontology Learning from Interactions in Natural Language and Reasoning. In *Proceedings of 24th Annual International Conference on Computer Science and Software Engineering* (pp. 120–132). Riverton, NJ, USA: IBM Corp. Retrieved from http://dl.acm.org/citation.cfm?id=2735522.2735536

Dinh, D., Reis, J. C. D., Pruski, C., Silveira, M. D., & Reynaud-Delaître, C. (2014). Identifying Change Patterns of Concept Attributes in Ontology Evolution. In V. Presutti, C. d'Amato, F. Gandon, M. d'Aquin, S. Staab, & A. Tordai (Eds.), *The Semantic Web: Trends and Challenges* (pp. 768–783). Springer International Publishing. doi:10.1007/978-3-319-07443-6_51

Dolbey, A., Ellsworth, M., & Scheffczyk, J. (2006). BioFrameNet: A Domain-specific FrameNet Extension with Links to Biomedical Ontologies. In *Proceedings of the "Biomedical Ontology in Action" Workshop at KR-MED* (pp. 87–94).

Duque-Ramos, A., Fernández-Breis, J. T., Iniesta, M., Dumontier, M., Egaña Aranguren, M., & Schulz, S. … Stevens, R. (2013). Evaluation of the OQuaRE framework for ontology quality. *Expert Systems with Applications, 40*(7), 2696–2703. doi:10.1016/j.eswa.2012.11.004

Duque-Ramos, A., Fernandez-Breis, J. T., Stevens, R., & Aussenac-Gilles, N. (2011). OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies. *Journal of Research and Practice in Information Technology*, *43*(2), 159.

Friedman, C., Kra, P., & Rzhetsky, A. (2002). Two biomedical sublanguages: A description based on the theories of Zellig Harris. *Journal of Biomedical Informatics*, *35*(4), 222–235. doi:10.1016/S1532-0464(03)00012-1 PMID:12755517

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gil, R., & Martin-Bautista, M. J. (2014). SMOL: A systemic methodology for ontology learning from heterogeneous sources. *Journal of Intelligent Information Systems*, *42*(3), 415–455. doi:10.100710844-013-0296-x

Grefenstette, G. (2012). *Explorations in Automatic Thesaurus Discovery*. Springer Science & Business Media.

Guo, T., Schwartz, D. G., Burstein, F., & Linger, H. (2015). Codifying Collaborative Knowledge: Using Wikipedia as a Basis for Automated Ontology Learning. In J. S. Edwards (Ed.), *The Essentials of Knowledge Management* (pp. 289–310). Palgrave Macmillan, UK. doi:10.1057/9781137552105_13

Hartung, M., Kirsten, T., Gross, A., & Rahm, E. (2009). OnEX: Exploring changes in life science ontologies. *BMC Bioinformatics*, *10*(1), 250. doi:10.1186/1471-2105-10-250 PMID:19678926

Jiang, X., & Tan, A.-H. (2010). CRCTOL: A Semantic-based Domain Ontology Learning System. *Journal of the American Society for Information Science and Technology*, *61*(1), 150–168. doi:10.1002/asi.21231

Klein, M., Fensel, D., Kiryakov, A., & Ognyanoff, D. (2002). Ontoview: Comparing and versioning ontologies. In Collected Posters ISWC 2002. Sardinia, Italy.

Maedche, A., Motik, B., Stojanovic, L., Studer, R., & Volz, R. (2002). Managing Multiple Ontologies and Ontology Evolution in Ontologging. In M. A. Musen, B. Neumann, & R. Studer (Eds.), *Intelligent Information Processing* (pp. 51–63). Springer, US. doi:10.1007/978-0-387-35602-0_6

Mayor, C., & Robinson, L. (2014). Ontological realism, concepts and classification in molecular biology: Development and application of the gene ontology. *The Journal of Documentation*, *70*(1), 173–193. doi:10.1108/JD-06-2013-0076

Moreda, P., Llorens, H., Saquete, E., & Palomar, M. (2011). Combining semantic information in question answering systems. *Information Processing & Management*, *47*(6), 870–885. doi:10.1016/j.ipm.2010.03.008

Ochoa, J. L., Martínez, J. M. R., Almela, Á., & Valencia-García, R. (2010). Efficient Multiword Term Extraction in Spanish. Application to the Financial Domain. In *Proceedings of ICIIT, 2010.* Lahore, Pakistan. Retrieved from https://www.academia.edu/928210/Efficient_Multiword_Term_Extraction_in_Spanish._Application_to_the_Financial_Domain

Ottens, K., Hernandez, N., Gleizes, M.-P., & Aussenac-Gilles, N. (2009). A Multi-Agent System for Dynamic Ontologies. *Journal of Logic and Computation*, *19*(5), 831–858. doi:10.1093/logcom/exn050

Paredes-Valverde, M. A., Valencia-García, R., Rodríguez-García, M. Á., Colomo-Palacios, R., & Alor-Hernández, G. (2015). A semantic-based approach for querying linked data using natural language. *Journal of Information Science*. doi:10.1177/0165551515616311

Petrucci, G., Ghidini, C., & Rospocher, M. (2016). Ontology Learning in the Deep. In E. Blomqvist, P. Ciancarini, F. Poggi, & F. Vitali (Eds.), *Knowledge Engineering and Knowledge Management* (pp. 480–495). Springer International Publishing. doi:10.1007/978-3-319-49004-5_31

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rodríguez-García, M. Á., Medina-Moreira, J., Lagos-Ortiz, K., Luna-Aveiga, H., García-Sánchez, F., & Valencia-García, R. (2016). Ontology-Based Platform for Conceptual Guided Dataset Analysis. In S. Omatu, A. Semalat, G. Bocewicz, P. Sitek, I. E. Nielsen, J. A. G. García, & J. Bajo (Eds.), *Distributed Computing and Artificial Intelligence, 13th International Conference* (pp. 155–163). Springer International Publishing. doi:10.1007/978-3-319-40162-1_17

Rodríguez-García, M. Á., Valencia-García, R., García-Sánchez, F., & Samper-Zapater, J. J. (2014a). Creating a semantically-enhanced cloud services environment through ontology evolution. *Future Generation Computer Systems*, *32*, 295–306. doi:10.1016/j.future.2013.08.003

Rodríguez-García, M. Á., Valencia-García, R., García-Sánchez, F., & Samper-Zapater, J. J. (2014b). Ontology-based annotation and retrieval of services in the cloud. *Knowledge-Based Systems*, *56*, 15–25. doi:10.1016/j.knosys.2013.10.006

Rubin, D. L., Shah, N. H., & Noy, N. F. (2008). Biomedical ontologies: a functional perspective. In *Presented at the 10th European Conference on Machine Learning (ECML 98) -- Workshop on Text Mining* (Vol. 9). Retrieved from http://bib.oxfordjournals.org/content/9/1/75

Ruiz-Martínez, J. M., Valencia-García, R., Martínez-Béjar, R., & Hoffmann, A. (2012). BioOntoVerb: A top level ontology based framework to populate biomedical ontologies from texts. *Knowledge-Based Systems*, *36*, 68–80. doi:10.1016/j.knosys.2012.06.002

Salas-Zárate, M. del P., Valencia-García, R., Ruiz-Martínez, A., & Colomo-Palacios, R. (2016). Feature-based opinion mining in financial news: An ontology-driven approach. *Journal of Information Science*. doi:10.1177/0165551516645528 doi:10.1177/0165551516645528

Sánchez, D., & Moreno, A. (2008). Learning non-taxonomic relationships from web documents for domain ontology construction. *Data & Knowledge Engineering*, *64*(3), 600–623. doi:10.1016/j.datak.2007.10.001

Sellami, Z., Camps, V., & Aussenac-Gilles, N. (2013). DYNAMO-MAS: A Multi-Agent System for Ontology Evolution from Text. *Journal on Data Semantics*, *2*(2–3), 145–161. doi:10.100713740-013-0025-1

Sinha, B., Garg, M., & Chandra, S. (2016). Identification and classification of relations for Indian languages using machine learning approaches for developing a domain specific ontology. In *Proceedings of the 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)* (pp. 415–420). doi:10.1109/ICCTICT.2016.7514617

Stojanovic, L., Maedche, A., Motik, B., & Stojanovic, N. (2002). User-Driven Ontology Evolution Management. In A. Gómez-Pérez & V. R. Benjamins (Eds.), *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web* (pp. 285–300). Springer Berlin Heidelberg. doi:10.1007/3-540-45810-7_27

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, *25*(1), 161–197. doi:10.1016/S0169-023X(97)00056-6

Terrientes, L. D. V., Moreno, A., & Sánchez, D. (2010). Discovery of Relation Axioms from the Web. In Y. Bi & M.-A. Williams (Eds.), *Knowledge Science, Engineering and Management* (pp. 222–233). Springer Berlin Heidelberg. doi:10.1007/978-3-642-15280-1_22

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tsai, R. T.-H., Chou, W.-C., Su, Y.-S., Lin, Y.-C., Sung, C.-L., & Dai, H.-J. … Hsu, W.-L. (2007). BIOSMILE: A semantic role labeling system for biomedical verbs using a maximum-entropy model with automatically generated template features. *BMC Bioinformatics, 8*, 325. doi:10.1186/1471-2105-8-325

Tsuruoka, Y., Tateishi, Y., Kim, J.-D., Ohta, T., McNaught, J., Ananiadou, S., & Tsujii, J. (2005). Developing a Robust Part-of-Speech Tagger for Biomedical Text. In P. Bozanis & E. N. Houstis (Eds.), *Advances in Informatics* (pp. 382–392). Springer Berlin Heidelberg. doi:10.1007/11573036_36

United States National Library of Medicine. (1996). PubMed. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed

Vila, K., Mazon, J.-N., & Ferrandez, A. (2011). Model-driven development for adapting question answering systems to restricted domains. *Journal of Research and Practice in Information Technology*, *43*(2), 141.

Wattarujeekrit, T., Shah, P. K., & Collier, N. (2004). PASBio: Predicate-argument structures for event extraction in molecular biology. *BMC Bioinformatics*, *5*(1), 155. doi:10.1186/1471-2105-5-155 PMID:15494078

Zablith, F., Sabou, M., d'Aquin, M., & Motta, E. (2009). Ontology Evolution with Evolva. In L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, … E. Simperl (Eds.), *The Semantic Web: Research and Applications* (pp. 908–912). Springer Berlin Heidelberg. doi:10.1007/978-3-642-02121-3_80

This research was previously published in the Journal of Information Technology Research (JITR), 10(4); edited by Francisco José García-Peñalvo; pages 56-69; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 54
Utilizing Combined Claims and Clinical Datasets for Research Among Potential Cases of Rare Diseases

Kevin J. Bennett
University of South Carolina, USA

Joshua Mann
University of Mississippi Medical Center, USA

Lijing Ouyang
Centers for Disease Control & Prevention, USA

ABSTRACT

With data quality issues with administrative claims and medically derived datasets, a dataset derived from a combination of sources may be more effective for research. The purposes of this article is to link an EMR-based data warehouse with state administrative data to study individuals with rare diseases; to describe and compare their characteristics; and to explore research with the data. These methods included subjects with diagnosis codes for one of three rare diseases from the years 2009-2014; Spina Bifida, Muscular Dystrophy, and Fragile X Syndrome. The results from the combined data provides additional information that each dataset, by itself, would not contain. The simultaneous examination of data such as race/ethnicity, physician and other outpatient visit data, charges and payments, and overall utilization was possible in the combined dataset. It is also discussed that combining such datasets can be a useful tool for the study of populations with rare diseases.

DOI: 10.4018/978-1-7998-1204-3.ch054

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The growth in the use of health information technology (HIT) has been substantial in the United States of America. The percentage of office-based physicians who utilized some form of an electronic health record (EHR) increased from 18% in 2001 to 83% in 2014 (Hsiao & Hing, 2014). Likewise, 96% of US hospitals had an electronic health record system as of 2015, although this percentage was lower for smaller or rural hospitals (68 & 70%, respectively) (Hsiao & Hing, 2014). HIT utilization has also grown across health care sectors, such as in emergency departments (Selck & Decker, 2016) and home health agencies (Mitchell, Bennett, & Probst, 2013). Much of this growth was spurred by the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act, which gave financial incentives to institutions that adopted electronic medical record (EMR) and other HIT technologies. The HITECH Act introduced meaningful use (MU), which established a set of criteria to ensure that the HIT adopted was clinically useful and eligible for the incentives (Blumenthal & Tavenner, 2010). Part of the MU criteria is to view, download, or transmit patient data electronically (health information exchange – HIE); as of 2015, 33% of providers and 14% of hospitals were able to do this (ONC, DHHS). In addition, up to 80% of hospitals were able to conduct some form of external electronic query of patient data (ONC, DHHS).

Health information also holds promise for health services researchers. By capturing patient, visit, and other data during encounters, researchers are able to get a fuller picture of health care utilization, outcomes, and the factors that affect them both. Data derived from clinically-based HIT systems is limited in its utility for researchers in several ways. First, data can be difficult to obtain unless the researcher has an established relationship with the HIT system owner (i.e., providers or organizations) and a method to retrieve the desired data. Additional difficulties can arise due to limitations or exclusions from the data to protect privacy (ensuring all data is adequately protected and/or de-identified), obtaining patient consent for the use of the data, or ensuring the completeness and appropriateness of the data extracted. Second, there are few HIT-based data sources that are representative of larger populations beyond service areas or insured populations. While larger hospital systems may have a larger service footprint, their data may still only capture their patient population and not be representative of an entire community, or may only capture one segment of care within the care continuum, such as inpatient care or ambulatory care. Third, HIT systems often have separate service lines, meaning that one department may not use the same HIT system or dataset as another department in that same health care system or network. While some institutions are working towards unified HIT systems and datasets (e.g., Veterans Administration, Kaiser Permanente), these are also limited to data from their own enrollees. Finally, data is often collected and catalogued in a manner that serves the clinical care needs of providers and delivery systems, and not the needs of researchers (Miriovsky, Shulman, & Abernethy, 2012; Rosenbloom et al., 2011).

For these reasons, the Healthcare Information and Management Systems Society includes in its definition of an electronic health record (EHR) the need to support ongoing performance measurement as well as support clinical research and evidence based research, facilitated by health information exchange (Handler et al., 2016). Larger databases that include clinical data derived from electronic health records have potential to provide sources of research data, while also providing an opportunity to shift the thinking of how research is conducted, towards a more rapid-based learning model (Abernethy et al., 2010; Etheredge, 2007). This model seeks to capitalize on the clinical content of the data, but the regular updating of such data allows for more time sensitive approaches to research.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

There is abundant research performed using a variety of data sources, including insurance claims data, surveys, and vital records (i.e., birth and death certificates). These data are utilized to examine a variety of topics, and have resulted in a wide range of findings (Büchele et al., 2016; Faurot et al., 2015; Lauer & McCallion, 2015; Martin, Osterman, & Sutton, 2010; Ross et al., 2000); however, they can be limited. Claims data are useful repositories of individual and organizational information, but often lack key patient and clinical information (such as comorbidities, severity of the diagnoses, lab values, or patient preferences). Other data sources may include more in-depth patient information, but may be based upon patient self-report of conditions, utilization, or health outcomes and therefore may not be reliable (National Center for Health Statistics, 2015; Pierannunzi, Hu, & Balluz, 2013). The ability to utilize data drawn from clinical documentation would be helpful for improving the quality, depth, and explanatory power for health services research by supplementing these administrative sources.

In 2014, a project began that sought to utilize a combined dataset derived from a claims-based dataset and a medical record based dataset for the years 2009-2014. The purpose of this project, described below, was to determine how to link a clinical dataset with state administrative and claims datasets to analyze data on people with potential cases of rare diseases (e.g., Spina Bifida, Muscular Dystrophy, and Fragile X Syndrome). We then sought to describe and compare the characteristics of the individuals identified, and to explore potential research questions and activities that can be conducted using combined data.

METHODS

Dataset Creation

This particular project, funded by the Disability Research and Dissemination Center, focused on individuals with billing codes for one of three rare diseases: Spina Bifida (International Classification of Diseases version 9 (ICD9) 741.xx), Muscular Dystrophy (ICD9 359.0, 359.1, and 359.21), and Fragile X Syndrome (ICD9 759.83). These conditions were chosen for study as they are relatively rare in the general population, and are therefore not widely studied. Individuals were included in each dataset if they had at least one encounter with a diagnosis code including one of these codes. The data are limited, however, in the ability to confirm a medical diagnosis of such diseases without full medical chart review.

Claims and Administrative Data

The Revenue and Fiscal Affairs (RFA) Office[1] is a neutral service entity that warehouses various administrative data from state agencies, which is able to link persons across multiple service providers, locations, and data sources. Through a series of statutes and agreements, state agencies and organizations grant the RFA access to data and data systems. For this project, the RFA provided two sources of data; the first was from Medicaid and State Health Plan Claims, and the second was from Hospital administrative datasets (all described below). The RFA has developed a series of algorithms using combinations of personal identifiers to create its own unique identifier, which enables analysts to link data by person across the multiple sources, while protecting confidentiality of the patients. All data requests were approved by all participating agencies and organizations, including the South Carolina Data Oversight Council and the University of South Carolina Institutional Review Board.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The first dataset included claims for patients with a service coded with one of the above mentioned rare disease billing codes from State Health Plan (SHP) or Medicaid (fee for service) as their insurance provider (hereafter referred to as the SHP/Medicaid dataset).

The SHP is an insurance option, administered by Blue Cross / Blue Shield of South Carolina that is offered to state and local government employees and their dependents. The other dataset is based upon information from the UB-04 billing and HCFA-1500 claim forms, utilized by facilities to bill for their services. These data include diagnostic information (International Statistical Classification of Diseases version 9, ICD-9), types of services received (either ICD-9, Healthcare Common Procedure Coding System(HCPCS), or Current Procedural Terminology (CPT) codes), service dates (admission and/or discharge dates), patient disposition, and other visit information for all inpatient and emergency department visits in the state, regardless of payer (hereafter referred to as the all-payer dataset).

The RFA queried these two data sources for patients with the ICD9 codes as above, and provided two separate datasets (one for SHP/Medicaid, and the other for the all-payer claims data). Race/ethnicity is not included in the State Health Plan data, and is thus not included in the SHP/Medicaid dataset; other information included in this dataset were age group, gender, and information regarding encounters and medications filled.

Electronic Health Records Data

SC is also home to Health Sciences South Carolina (HSSC), which houses the Clinical Data Warehouse (CDW)[2]. HSSC is a statewide research collaborative focused on improving the quality of healthcare delivery, improving health information systems, training the healthcare workforce, and improving patient outcomes. Partners include seven of the state's largest health systems and the state's largest research-intensive universities. This is among the first systems to bring together data from three major research universities and several large health care systems. HSSC created the CDW to facilitate these goals; the CDW aggregates de-identified electronic patient records from these systems to enable providers and researchers to access patient data for research or treatment purposes. The data collected are from electronic health records related to inpatient, outpatient, and emergency department encounters from the affiliated systems.

Using the ICD9 codes described above, individuals were identified as potentially having one of these conditions in the CDW dataset (2009-2014). Information extracted from this dataset included patient demographics (sex, race, ethnicity, and area of residence – rural or urban), visit information (date, visit type (emergency, inpatient, other)), payer (commercial, government, other, self-pay), admission and discharge information, and diagnoses associated with the visit.

HSSC then sent their data file to the analysts at the RFA to merge with the two state-based datasets. To further alleviate privacy concerns, the service dates for all data were shifted randomly, but by a fixed amount across an individual's encounters. The RFA then utilized a proprietary linkage algorithm to match patients from the three sources (CDW, Medicaid/SHP, all-payer datasets). Each person with unduplicated data was assigned a random number generated by an algorithm, referred to as the Unique ID. The algorithm uses personal identifiers that include, but are not limited to: first name, middle initial, last name, date of birth, race, and gender. The dataset was edited (i.e., characters were removed from numeric fields, dates were compared to valid ranges) and standardized (i.e., all characters were converted to uppercase) before being run through the algorithm. All of the identifier fields on a record housed at RFA were compared to records in the CDW data to identify exact or partial matches of individuals. A

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

'MatchType' (a string of characters that represent the match for each identifier field) was then assigned for each potential match. A 'MatchScore' (the rate indicating how often this 'MatchType' yields the same person when compared across multiple rows of data) and a 'FieldScore' (the number indicating the number of the fields that matched) was then assigned to the record. For each individual original record for which there were records with a MatchScore of 90 or greater, a single record was chosen as the actual match based on a match with more specific identifiable fields, such as social security number and date of birth. Once that single record match was chosen, the unique ID number from the matched record was assigned. If no records for a given original record have a MatchScore of 90 or greater, candidates for that original record were then compared to other possible candidates with lower MatchScore values. The process of scoring was then repeated for all records in this pool. The algorithm accounts for misspelling, name changes, and slight differences in the date of birth. An individual's Unique ID remains associated with that individual on all subsequent episodes of services in data captured by RFA, regardless of data source or service provider. Those that had an initial MatchScore of at least 90 were considered strong matches. The remaining were then compared on other characteristics to find additional matches, with a lower limit MatchScore of 60.

The final merged dataset was created using the RFA-generated unique identifier. The demographic variables were then compared and reconciled. In cases of a discrepancy, the assignment was based upon one of the state based datasets (all-payer or SHP/Medicaid).

Health care encounters were matched based upon the patient's ID, the data of the visit, and the setting and categorized into five categories: Inpatient, Emergency Department, Outpatient, Office Visits, and Other. Each dataset included different taxonomies to label their encounters; each utilized Inpatient, Emergency department, and Outpatient while the SHP/Medicaid included both Office and Other categories. To avoid double counting a visit recorded in more than one data source, a stratified matching method was utilized. First, encounters with the same date, by ID, were identified. Next, depending upon the dataset, the category of service was identified (Inpatient, Outpatient, or Emergency Department). Those labeled as 'Office" encounters in the SHP/Medicaid were matched with encounters labelled as "Outpatient" in the CDW data when appropriate. Second, if the date and category matched on two encounters, this was determined to be one visit. Non-matches were designated as different encounters.

Merging the data is simple enough, if done by the unique identifier identified by the matching algorithm described above. In practice, a major hurdle to an accurate merge was in identifying a single encounter as just one encounter, and not confusing the multiple claims associated with an encounter as separate events. For example, a single physician visit may result in multiple claims, associated with tests, labs, or consultations. Multiple claims were merged into a single event for the accuracy of counting encounters. Also, with the merge, the same event could be represented in each of the datasets; thus, care must be taken to match events by date of service, service type, and service location. In addition, professional claims were also added to applicable events (e.g., professional claims associated with an inpatient claim or outpatient surgery).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

RESULTS

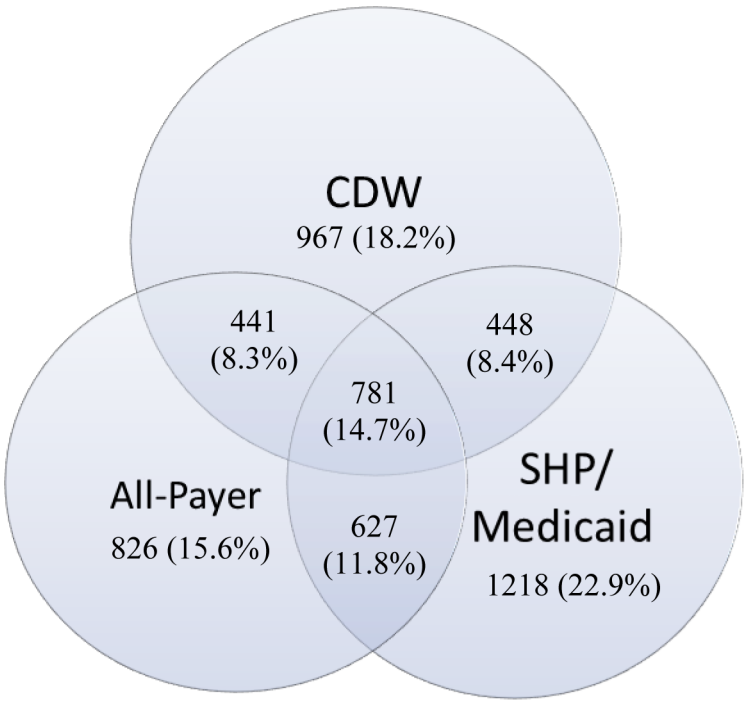
Description of the Data

The resulting dataset derived from the above process was comprised of 2,637 patients identified within the CDW; 3,074 within the SHP/Medicaid data; and 2,675 from the all-payer data, for a total of 5,308 unique patients. When merged, 8.4% were in both the CDW and SHP/Medicaid; 8.3% were in both the CDW and all-payer; 11.8% were in both the SHP/Medicaid and all-payer; and 14.7% were in all three. Thus, 56.7% of the subjects were found in only one of the three datasets (see Figure 1).

Table 1 displays the demographic characteristics of those in each dataset, compared to those in all three datasets. The all-payer data had a higher proportion of females (56.2%) than the other datasets; the combined data was comprised of 54.2% female. The all-payer data also had the highest proportion that were white (72.9%), compared to 68.1% for the CDW dataset and 54.5% for the combined data (SHP/Medicaid race data were suppressed). The all-payer data skewed older but the combined data skewed towards the 0-17 age group (39.1%).

Table 2 displays the demographic characteristics of the study population by disease. Of these, 61.2% had a code associated with Spina Bifida, 32.7% had a code associated with Muscular Dystrophy, and 6.1% had a code associated with Fragile X. Potential Spina Bifida patients were predominantly female, white, and under the age of 17. Potential Muscular dystrophy patients were predominately male, white, and under the age of 17. Potential Fragile X patients were also predominately male, white, and under the age of 17.

Figure 1. Resulting dataset comprised of 2,637 patients identified within the CDW



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Demographic characteristics of patients in administrative, clinical, and combined datasets; South Carolina, 2009-2014

	CDW†		SHP/Medicaid†		All-Payer†		All 3, Combined	
	N	%	N	%	N	%	N	%
	2637		3074		2675		5308	
Sex								
Female	1379	52.3	1608	52.3	1503	56.2	2875	54.2
Male	1212	46.0	1466	47.7	1172	43.8	2419	45.6
Missing/Unknown	46	1.7	0	0.0	0	0.0	14	0.3
Race/Ethnicity*								
White	1796	68.1			1950	72.9	2895	54.5
African American	621	23.5			584	21.8	913	17.2
Hispanic	110	4.2			63	2.4	131	2.5
Other	110	4.2			78	2.9	151	2.8
Missing/Unknown	0	0.0			0	0.0	1218	22.9
Age Group								
0-17	1016	38.5	1568	51.0	715	26.7	2078	39.1
18-24	248	9.4	283	9.2	224	8.4	417	7.9
25-44	618	23.4	796	25.9	863	32.3	1404	26.5
45-64	510	19.3	364	11.8	612	22.9	992	18.7
65+	245	9.3	63	2.0	261	9.8	417	7.9

*Race/ethnicity was suppressed from the SHP/Medicaid Dataset

†*Proportions differ significantly, $\alpha = 0.05$, by data source*

Utilization and Expenditures

Table 3 displays the summary of annual claims, by type and year, for the individuals in each dataset. For the combined data, 22.7% had one or more inpatient visits. This proportion varied by condition; a smaller proportion of those potentially with Fragile X had at least 1 inpatient stay (13.0%) compared to those potentially with Muscular Dystrophy (22.6%) and Spina Bifida (23.8%). This trend remained, with some variations, when annualized (see Table 3). For emergency department visits, 38.3% of the entire sample had one or more visits. This proportion also differed by condition; a smaller proportion of those potentially with Fragile X had one or more visit (26.2%) compared to those potentially with Muscular Dystrophy (36.5%) and Spina Bifida (40.5%). This trend also remained, with some variations, when annualized (see Table 3).

We next examined median annual charges, subset by category of service. For individuals with potential cases of any of the three conditions, combined inpatient visits had an annual median charge of $25,410 and ED visits had an annual median charge of $13,067. Inpatient charges were lower for potential Muscular Dystrophy and Fragile X patients, but higher for potential Spina Bifida patients. The ED charges were similar for Spina Bifida, slightly higher for Muscular Dystrophy, but lower for potential Fragile X patients.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Demographic characteristics of patients potentially with rare diseases in a combined administrative and clinical dataset; South Carolina, 2009-2014

	Spina Bifida†		Muscular Dystrophy†		Fragile X†		Any of the Three	
	N	%	N	%	N	%	N	%
All	3252	61.2	1737	32.7	319	6.1	5308	100
Sex								
Female	1990	61.2	795	45.8	90	28.2	2875	54.2
Male	1258	38.7	932	53.7	229	71.8	2419	45.6
Missing/Unknown	4	0.1	10	0.6	0	0.0	14	0.3
Race/Ethnicity								
White	1707	52.5	1053	60.6	135	42.3	2895	54.5
African American	586	18.0	271	15.6	56	17.6	913	17.2
Hispanic	80	2.5	44	2.5	7	2.2	131	2.5
Other	91	2.8	45	2.6	15	4.7	151	2.8
Missing/Unknown	788	24.2	324	18.7	106	33.2	1218	22.9
Age Group								
0-17	1382	42.5	505	29.1	191	59.9	2078	39.1
18-24	301	9.3	87	5.0	29	9.1	417	7.9
25-44	948	29.2	397	22.9	59	18.5	1404	26.5
45-64	482	14.8	482	27.7	28	8.8	992	18.7
65+	139	4.3	266	15.3	12	3.8	417	7.9

†Proportions differ significantly, $\alpha = 0.05$, by condition

An examination of the primary diagnoses for type of visit was also instructive (see Table 4). For ED visits, commonly cited diagnoses included upper respiratory infections, urinary tract disorders, digestive system disorders, and skin ailments. In addition to those listed above, several diagnoses were unique to the condition: for potential Spina Bifida patients, migraine headaches were the fourth most common primary diagnosis; for potential Muscular Dystrophy patients, schizophrenic disorders were the sixth most common; for potential Fragile X patients, epilepsy was the fourth most common.

For inpatient visits, commonly cited primary diagnoses included urinary tract disorders, live births, pneumonia, and complications from procedures. In addition to those listed above, several diagnoses were unique to the condition: for potential Spina Bifida patients, chronic skin ulcers were the second most common primary diagnosis; for potential Muscular Dystrophy patients, cellulitis and abscesses was the tenth most common; for potential Fragile X patients, Bulbus cordis anomalies and anomalies of cardiac septal closure were the most common primary diagnosis.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Percent distribution of annual visits by patients potentially with rare diseases, by type of encounter, and year, in the combined administrative and clinical dataset; South Carolina, 2009-2014

# of Visits	Spina Bifida†			Muscular Dystrophy†			Fragile X†			Any of the Three		
	0	1	>1	0	1	>1	0	1	>1	0	1	>1
Inpatient												
2009	77.3	14.4	8.3	78.1	13.0	8.8	88.8	5.2	6.0	78.3	13.4	8.3
2010	74.2	15.6	10.2	76.2	14.3	9.4	86.6	7.4	5.9	75.6	14.7	9.7
2011	74.0	14.8	11.2	73.6	14.7	11.7	83.7	10.0	6.3	74.5	14.5	11.1
2012	74.3	15.5	10.2	77.5	12.3	10.3	84.4	8.1	7.4	76.0	14.0	10.1
2013	75.6	14.8	9.6	76.0	13.5	10.6	86.7	6.8	6.5	76.4	13.9	9.8
2014*	83.9	11.7	4.4	85.6	10.3	4.0	92.6	7.4‡		85.0	10.8	4.2
Emergency Department												
2009	57.8	21.7	20.5	61.4	22.6	16.0	69.1	18.9	12.0	59.8	21.8	18.4
2010	56.1	22.8	21.1	61.5	20.7	17.7	75.5	13.0	11.5	59.2	21.5	19.4
2011	56.8	22.1	21.2	60.4	20.3	19.3	65.9	21.9	12.2	58.6	21.5	20.0
2012	56.4	21.5	22.1	60.2	20.6	19.3	71.5	15.6	13.0	58.6	20.8	20.6
2013	57.2	21.3	21.6	61.7	19.9	18.3	73.8	14.7	11.5	59.7	20.4	19.9
2014*	76.5	16.1	7.5	80.6	13.3	6.1	88.6	7.9	3.5	78.6	14.7	6.8

†*Proportions differ significantly, α = 0.05, by year and service category*

‡ *Categories combined due to small sample sizes*

*Through August 2014

DISCUSSION

Using administrative claims and clinical data allowed for a more comprehensive analysis of the experiences of patients with potential cases of selected rare diseases. Each of the datasets used in this project has its relative strengths; the CDW is derived from clinical documentation, the SHP/Medicaid data represents all of the individual level claims for those enrolled in those programs, and the all-payer data represents data from all facilities. In this study, the study population nearly doubled in size due to this combination compared to using any of the datasets alone. In addition, the combined dataset allowed researchers to obtain more comprehensive information about a continuum of care for patients across facilities and provider types. This is particularly useful for those with potential cases of rare diseases, as more complete information about their utilization can now be analyzed.

The use of combined administrative and clinical datasets may have the potential to improve studies of other populations as well, in the following ways:

- Matching visit data across the datasets can help fill in 'missing' information (such as race/ethnicity, visit type, location, prescriptions, lab values);
- Combining facility and professional claims can allow researchers to obtain more complete information for inpatient facility encounters;
- Examining actual laboratory values, not just that the test was performed, might permit researchers to study the linkage between such values and outcomes and/or utilization;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

*Table 4. Ranking and frequency of diagnoses, by patients potentially with rare diseases, by type of encounter, in the combined administrative and clinical dataset; South Carolina, 2009-2014**

	Spina Bifida (n)	Muscular Dystrophy (n)	Fragile X (n)
Emergency Department			
1	Chronic ulcer of skin (1103)	General symptoms (238)	General symptoms (36)
2	Symptoms involving head and neck (722)	Symptoms involving respiratory system / chest (215)	Acute upper respiratory infections of multiple sites (32)
3	Other disorders of urethra and urinary tract (682)	Other symptoms involving abdomen and pelvis (184)	Attention to artificial openings (29)
4	Migraine (646)	Other disorders of urethra and urinary tract (180)	Epilepsy and recurrent seizures (23)
5	Other symptoms involving abdomen / pelvis (497)	Symptoms involving digestive system (172)	Other congenital anomalies of upper alimentary tract (22)
6	Symptoms involving respiratory system / chest (457)	Schizophrenic disorders (140)	Diseases of hard tissues of teeth (20)
7	Other and unspecified disorders of back (446)	Other and unspecified disorders of back (126)	Other symptoms involving abdomen and pelvis (19)
8	General symptoms (403)	Diabetes mellitus (110)	Acute pharyngitis (16)
9	Symptoms involving digestive system (308)	Pneumonia, organism unspecified (109)	Diseases of esophagus (16)
10	Other disorders of bladder (339)	Chronic ulcer of skin (106)	Cellulitis and abscess of face (16)
Inpatient†			
1	Complications peculiar to certain specified procedures (212)	Care involving use of rehabilitation procedures (79)	
2	Chronic ulcer of skin (125)	Other diseases of lung (54)	
3	Other disorders of urethra and urinary tract (123)	Pneumonia, organism unspecified (52)	
4	Spina bifida (119)	Single liveborn (42)	
5	Other conditions of brain (100)	General symptoms (37)	
6	Single liveborn (98)	Muscular dystrophies and other myopathies (35)	
7	Other congenital anomalies of nervous system (84)	Other disorders of urethra and urinary tract (32)	
8	Other cellulitis and abscess (82)	Symptoms involving respiratory system / chest (28)	
9	Septicemia (73)	Complications peculiar to certain procedures (27)	
10	General symptoms (72)	Other cellulitis and abscess (26)	

*Diagnostic categories using ICD-9 codes and the Clinical Classification Software (CCS) system developed by the Healthcare Cost and Utilization Project: https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. †Inpatient diagnoses are not displayed for Fragile X due to small cell sizes, to ensure privacy

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Expanding information regarding insurance type could lead to more nuanced examination by payer status;
- Providing a more detailed description of the service unit within an inpatient facility can generate a better understanding of the care delivery context for the researcher;
- Identifying comorbidities or complexities in care due to the addition of diagnosis codes and procedure codes from the merged data might improve ability of researcher to make connections;
- Examining the provider classification, such as specialty, associated with the claims could allow for a more comprehensive understanding of what providers these individuals are seeking care from;
- Examining charge data associated with the claim can create a better understanding of health care costs associated with such diseases together with cost-to-charge ratios;
- Utilizing a combined dataset may be helpful to policy makers who are tasked with monitoring health services programs, interventions, and funding as their decisions can be made based upon fuller and more complete information as well as better examinations of the outcomes of such decisions.

There are limitations to these data, however. The CDW data included a large amount of outpatient visit data, but the categorization was vague, making it difficult. Specifically, it is not always possible to differentiate outpatient encounters by type or location (e.g., outpatient surgery vs. ambulatory physician visit). This can be clarified by matching to the SHP/Medicaid visit data, but only for those individuals with information in both sources. If these outpatient claims could be more precisely identified, the utility of that data may improve. Issues regarding the quality of clinical data are not uncommon, and continue to be a barrier to using such data for research (Coorevits et al., 2013; Weiskopf & Weng, 2013).

There were also cases of contradictory information within matched individuals and claims. For example, approximately 1% of the demographic data did not match exactly (i.e., sex, age and race/ethnicity not matching upon merge). When such discrepancies occurred, we relied on data from a state-based source rather than the CDW data. There were also instances where potential matched claims had substantially different service unit categorizations; it was often unclear if these simply represented different claims on the same day, or a discrepancy in classification by the provider. These instances were treated, conservatively, as matches. In addition, there are various reasons why an individual would be presenting one or more of the databases, but not all there, such as seeking care in only one place of service not captured in all three datasets, not having insurance coverage, or even seeking care from out-of-state providers.

Finally, identifying potential cases using ICD9 coding is not perfect. It has been shown that accuracy of administrative data was low when using a straightforward criterion of a single ICD9 diagnosis code for muscular dystrophy. However, accuracy of finding true cases improved substantially with added information in the algorithm (Smith et al., 2017). Future studies may combine medical chart review with administrative data to address the questions of ICD-9 codes validation for rare conditions so that these administrative data sources can be better used for research on rare conditions.

Despite these limitations and difficulties, such a merged dataset represents a useful tool for researchers and policy makers alike. The use of such data has the potential to lead to improvements in population health and service delivery in many aspects.

This analysis was supported by a cooperative agreement from the Centers for Disease Control & Prevention through the Disability Research and Dissemination Center. RFA-R14-002, EMR Rare Conditions. Rare Condition research using a combined Administrative / Clinical Data warehouse.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DISCLAIMER

The findings in this manuscript are those of the authors and do not necessarily represent the views of the South Carolina Department of Health and Human Services, the South Carolina Public Employee Benefit Authority, the South Carolina Data Oversight Council or the South Carolina Revenue and Fiscal Affairs Office.

REFERENCES

Abernethy, A. P., Etheredge, L. M., Ganz, P. A., Wallace, P., German, R. R., Neti, C., ... Murphy, S. B. (2010). Rapid-Learning System for Cancer Care. *Journal of Clinical Oncology*, *28*(27), 4268–4274. doi:10.1200/JCO.2010.28.5478 PMID:20585094

Blumenthal, D., & Tavenner, M. (2010). The "Meaningful Use" Regulation for Electronic Health Records. *The New England Journal of Medicine*, *363*(6), 501–504. doi:10.1056/NEJMp1006114 PMID:20647183

Büchele, G., Becker, C., Cameron, I. D., Auer, R., Rothenbacher, D., & König, H. H. (2016). Fracture risk in people with developmental disabilities: Results of a large claims data analysis. *Osteoporosis International*, *24*, 1–7. PMID:27553444

Coorevits, P., Sundgren, M., Klein, G. O., Bahr, A., Claerhout, B., Daniel, C., ... Kalra, D. (2013). Electronic health records: New opportunities for clinical research. *Journal of Internal Medicine*, *274*(6), 547–560. doi:10.1111/joim.12119 PMID:23952476

Etheredge, L. M. (2007). A Rapid-Learning Health System. *Health Affairs*, *26*(2), w107–w118. doi:10.1377/hlthaff.26.2.w107 PMID:17259191

Faurot, K. R., Jonsson Funk, M., Pate, V., Brookhart, M. A., Patrick, A., Hanson, L. C., ... Stürmer, T. (2015). Using claims data to predict dependency in activities of daily living as a proxy for frailty. *Pharmacoepidemiology and Drug Safety*, *24*(1), 59–66. doi:10.1002/pds.3719 PMID:25335470

Handler, T., Holtmeier, R., Metzger, J., Overhage, M., Taylor, S., & Underwood, C. (2016). HIMSS electronic health record definitional model version 1.0. Retrieved from www.himss.org

Hsiao, C. J., & Hing, E. Use and Characteristics of Electronic Health Record Systems Among Office-based Physician Practices: United States, 2001–2013. (2014). NCHS Data Brief No. 143. Available from: http://www.cdc.gov/nchs/data/databriefs/db143.pdf

Lauer, E., & McCallion, P. (2015). Mortality of people with intellectual and developmental disabilities from select US state disability service systems and medical claims data. *Journal of Applied Research in Intellectual Disabilities*, *28*(5), 394–405. doi:10.1111/jar.12191 PMID:25994364

Martin, J.A., Osterman, M.J.K., & Sutton, P.D. (2010). Are preterm births on the decline in the United States? Recent data from the National Vital Statistics System. *NCHS Data Brief*, (39).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Miriovsky, B. J., Shulman, L. N., & Abernethy, A. P. (2012). Importance of Health Information Technology, Electronic Health Records, and Continuously Aggregating Data to Comparative Effectiveness Research and Learning Health Care. *Journal of Clinical Oncology*, *30*(34), 4243–4248. doi:10.1200/JCO.2012.42.8011 PMID:23071233

Mitchell, J., Bennett, K. J., & Probst, J. (2013). Organizational Factors Associated with Health Information Technology Adoption and Utilization Among Home Health / Hospice Agencies. *International Journal of Healthcare Information Systems and Informatics*, *6*(3), 46–59. doi:10.4018/jhisi.2011070104

Rosenbloom, S. T., Denny, J. C., Xu, H., Lorenzi, N., Stead, W. W., & Johnson, K. B. (2011). Data from clinical notes: A perspective on the tension between structure and flexible documentation. *Journal of the American Medical Informatics Association*, *18*(2), 181–186. doi:10.1136/jamia.2010.007237 PMID:21233086

Ross, N. A., Wolfson, M. C., Dunn, J. R., Berthelot, J.-M., Kaplan, G. A., & Lynch, J. W. (2000). Relation between income inequality and mortality in Canada and in the United States: Cross sectional assessment using census data and vital statistics. *BMJ (Clinical Research Ed.)*, *320*(7239), 898–902. doi:10.1136/bmj.320.7239.898 PMID:10741994

Selck, F.W., & Decker, S.L. (2016). Health Information Technology Adoption in the Emergency Department. *Health Serv Res.*, *5*1(1), 32–47.

Smith, M. G., Royer, J., Mann, J. R., & McDermott, S. (2017, January 12). Using Administrative Data to Ascertain True Cases of Muscular Dystrophy: Rare Disease Surveillance. *JMIR Public Health and Surveillance*, *3*(1), e2. doi:10.2196/publichealth.6720 PMID:28082256

The Office of the National Coordinator for Health Information Technology, U.S. Department of Health and Human Services. (2014). Electronic Health Information Exchange Performance Reported to the Medicare EHR Incentive Program. Retrieved from http://dashboard.healthit.gov/quickstats/pages/eligible-provider-electronic-hie-performance.php

The Office of the National Coordinator for Health Information Technology, U.S. Department of Health and Human Services. (2016). Hospital Progress to Meaningful Use by Size, Type, and Urban/Rural Location. Retrieved from http://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-Progress-to-Meaningful-Use-by-size-practice-setting-area-type.php

The Office of the National Coordinator for Health Information Technology, U.S. Department of Health and Human Services. (2016). Hospital Capability to Electronically Query. Retrieved from http://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-Electronic-Query-Capability.php

Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. *Journal of the American Medical Informatics Association*, *20*(1), 144–151. doi:10.1136/amiajnl-2011-000681 PMID:22733976

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] See http://rfa.sc.gov/healthcare

[2] See https://healthsciencessc.org/tools-services/tool/clinical-data-warehouse/219

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 13(2); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 1-12; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 55
PAGeneRN: Parallel Architecture for Gene Regulatory Network

Dina Elsayad
Ain Shams University, Egypt

A. Ali
Ain Shams University, Egypt

Howida A. Shedeed
Ain Shams University, Egypt

Mohamed F. Tolba
Ain Shams University, Egypt

ABSTRACT

The gene expression analysis is an important research area of Bioinformatics. The gene expression data analysis aims to understand the genes interacting phenomena, gene functionality and the genes mutations effect. The Gene regulatory network analysis is one of the gene expression data analysis tasks. Gene regulatory network aims to study the genes interactions topological organization. The regulatory network is critical for understanding the pathological phenotypes and the normal cell physiology. There are many researches that focus on gene regulatory network analysis but unfortunately some algorithms are affected by data size. Where, the algorithm runtime is proportional to the data size, therefore, some parallel algorithms are presented to enhance the algorithms runtime and efficiency. This work presents a background, mathematical models and comparisons about gene regulatory networks analysis different techniques. In addition, this work proposes Parallel Architecture for Gene Regulatory Network (PAGeneRN).

DOI: 10.4018/978-1-7998-1204-3.ch055

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Bioinformatics is the application of information technology in the field of molecular biology to manage process and analyze both genomic and molecular biological data. Its primary goal is to improve the understanding of biological processes (Nair, 2007). Bioinformatics is a very rich filed of research, in fact it includes so many research areas with diverse applications and according to Srikanth Aluru (2006), the major research areas in Bioinformatics are: Genome annotation (Kelley, MacCallum, & Sternberg, 2000), Sequence analysis (Durbin 1998, Sharif et al. 2015, Sharif et al. 2016, Tharwat et al. 2015), Computational evolutionary biology (Cosmides & Tooby, 1995), Gene expression analysis (Velculescu, Zhang, Vogelstein, & Kinzler, 1995), Analysis of regulation (Janssen, Jones, Jones, & Woods, 1988) and Protein expression analysis (Ghaemmaghami, et al., 2003). In addition, it includes analysis of mutations in cancer (Berrozpe, Schaeffer, Peinado, Real, & Perucho, 1994), Predictions of protein structure (Shortle, 2000), Comparative Genomics (Rubin, et al., 2000), Modeling biological systems (Haefner, 2005), High throughput image analysis (Dowsey, 2005), Protein-protein docking (Dominguez, Boelens, & Bonvin, 2003)and Microarrays (Churchill, 2002).

The microarray is one of the active research fields in bioinformatics. It's a multiplex lab-on-a-chip that assays large amounts of biological material using high throughput screening methods (Culf, Cuperlovic-Culf, & Ouellette, 2006). Moreover, the microarrays experiment (Li & Li, 2008) starts with the sample preparation after that the hybridization, diction step and finally the data acquisition and analysis step. The microarray chips are scanned using special devices called Microarray Scanner. The microarray type depends on the biological material that assays. There are a number of types of microarrays including DNA microarrays (Gasch, et al., 2000), MMChips, for surveillance of microRNA populations (Meenakshisundaram, et al., 2009), Protein microarrays (Stoevesandt, Taussig, & He, 2009), Tissue microarrays (Kononen, et al., 1998), Cellular microarrays (Ma & Horiuchi, 2006), Chemical compound microarrays (Ma & Horiuchi, 2006), Antibody microarrays (Rivas, et al., 2008) and Carbohydrate arrays (glycoarrays) (Culf, Cuperlovic-Culf, & Ouellette, 2006).

The DNA Microarray technology is a high throughput experimental technique that can measure expression levels of hundreds of thousands of genes simultaneously. As an illustration, expression level of the gene is estimated by measuring the amount of mRNA for that gene. A gene is active if it is being transcribed; the more mRNA usually indicates more gene activity (Schena, Shalon, Davis, & Brown, 1995). The DNA microarray has become a useful technique in gene expression analysis for the development of new diagnostic tools and for the identification of disease genes and therapeutic targets for disease such as cancer. In fact, it's the basis for a very efficient screening and diagnosis of disease in an early stage of development. Microarrays could also be expanded into a starting point for developing new treatments for various types of disease.

The analysis of the data resulting from microarrays remains a big challenge for the huge volume of data it produces. In addition, the data analysis process involves various computational tasks (Yang, et al., 2008) and one of the microarrays data analysis tasks is network analysis (Haman & Valenta, 2013). The gene regulatory network (Aluru, 2006)goal is to study the topological organization of the genes interactions. The gene regulatory network analysis techniques can be classified into four main categories: component analysis techniques (Raychaudhuri, Stuart, & Altman, 2000) (Watkins, 2004) (Holter, et al., 2000) (Hyvarinen, Karhunen, & Oja, 2001) (Aapo, 1999) (Liebermeister, 2002) (Liao, et al., 2003) (Chang, Ding, Hung, & Fung, 2008), mutual information based techniques (Romualdi & Chiara, 2011) (Shi, Schmidt, Liu, & Muller-Wittig, 2011), reverse engineering techniques (Jostins & Jaeger, 2010)[44] and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the regression techniques (Gregoretti, Belcastro, Di Bernardo, & Oliva, 2010). In fact, almost all of these techniques are computational intensive and time-consuming. Therefore, parallel algorithms are needed to enhance the techniques performance. The parallelism techniques are categorized as shared memory (Karp, 1988) (Hager & Wellein, 2010) and distributed memory (Hager & Wellein, 2010).Moreover, these two models can be mixed in the hybrid model to take the full advantage of the computing power (Hager & Wellein, 2010).Another form of the distributed computing is cloud computing (Erl, Puttini, & Mahmood, 2013). In addition, the graphic processing unit (GPU) can be used to enhance performance.

This work presents a background, mathematical models and comparisons about gene regulatory networks analysis different techniques. In addition, this work proposes a parallel architecture for gene regulatory network called PAGeneRN (**P**arallel **A**rchitecture for **Gene R**egulatory **N**etwork). This work is organized as follows; the related work that contains gene regulatory network analysis different techniques is discussed in the second. While, the third section provides information about the microarrays, microarrays experiments and the parallelism different techniques. The gene regulatory network analysis techniques are categorized into four categories; component analysis, mutual information, reverses engineering and finally regression techniques. These techniques challenges are summarized in the fourth section. While, the fifth section provides an architecture which enhances the gene regulatory network analysis technique. Besides that, the conclusion is showed in the last section.

RELATED WORK

One of the microarray data analysis techniques (will be discussed in the next section) is the network analysis (Haman & Valenta, 2013)which is a powerful analysis tool to understand real world phenomena and one type of the network analysis is the gene regulatory network, actually, the gene regulatory network focuses on studying gene interactions. This section illustrates the gene regulatory network analysis different techniques, in addition; the gene regulatory network analysis techniques are discussed from sequential and parallel points of view. Furthermore, the applications of gene regulatory network analysis on time series data are discussed.

Gene Regulatory Network

The gene regulatory network aims to study the genes interaction topological organization and gene-to-gene influences. As shown in Figure 1 the gene regulatory network consists of a set of nodes and sets of edges, where the nodes represent genes and edges represent regulatory interactions between genes.

The illustrated example represents a regulatory network in mouse **E**mbryonic **S**tem Cells (ESC) (Zhou, Chipperfield, Melton, & Wong, 2007) the master regulators are Oct4, Sox2 and Nano, where pink nodes represent the main regulators and their protein regulators are represented in yellow. Moreover, the regulatory interactions are illustrated by blue and pink arrows, while, the protein interactions are represented by the orange and green lines. In addition, the arrows from the dashed ellipse indicate that all the regulators inside the ellipse regulate the target.

As illustrated in Figure 2 the gene regulatory network analysis techniques can be classified into four categories: Component analysis techniques, mutual information based techniques, reverse engineering techniques and regression techniques. Furthermore, the component analysis technique can be further categorized into a number of sub-classes. As mentioned before the component analysis of microarrays

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A regulatory network in mouse ESC anchored[1]

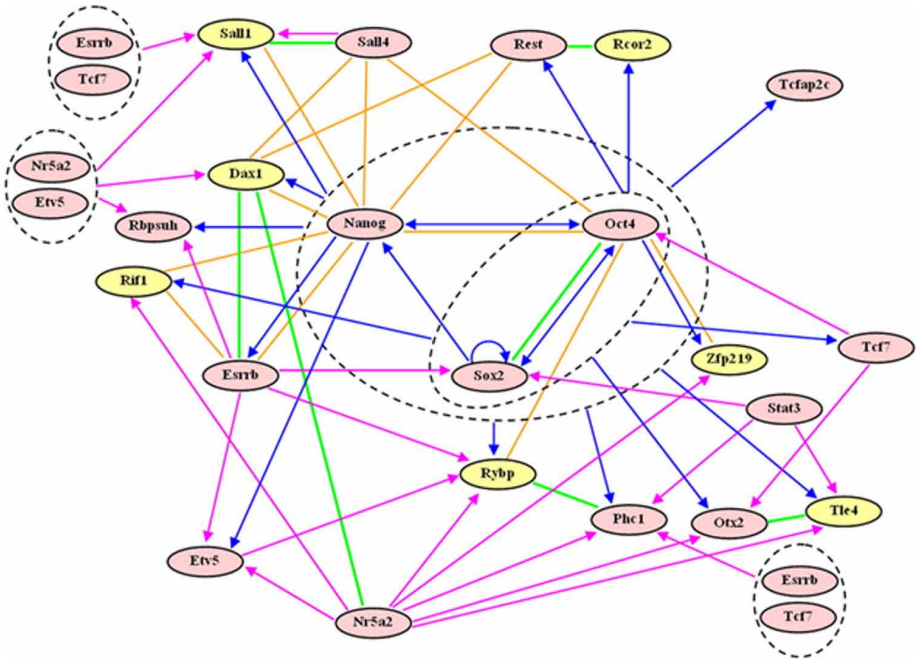
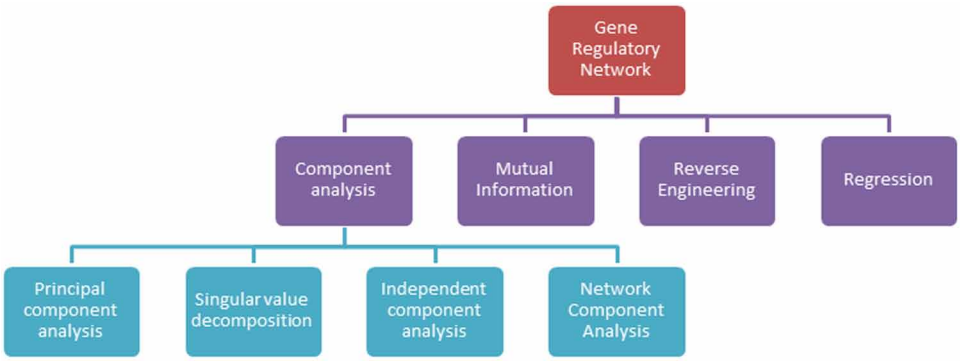


Figure 2. Gene regulatory network techniques



data aims at studying the ways in which gene responses vary under different conditions. While, here the component analysis is a computational technique and aims to build the gene regulatory network. One of the component analysis techniques is the principal component analysis technique (PCA) (Raychaudhuri, Stuart, & Altman, 2000).The PCA is the dimensionality reduction technique which determines the key variables that explain the difference in the data and simplify the data analysis and visualization. In other words, when applying PCA on a data matrix [*A*] of *M* observation in *N* variables; in the case of the gene expression data, the observation are the genes and the variables are the experiment conditions/ samples, it finds new *x* variables which reduce the dimensionality of [*A*], where, the new *x* variables must be orthogonal and mutually uncorrelated. Also, it must take into consideration covering as much as possible the variance in the original variables. These variables are called the principle components. The principle component is a linear combination of the original variable.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Applying the PCA on gene expression data allows scientists to summarize the ways in which gene responses vary under different conditions. When applying the PCA to gene expression data can consider the genes or samples or both as variables. In case of genes being considered as variables, a set of "principle gene components" are created. This principle gene component set illustrates the experimental responses produced by these genes. In the other hand, the PCA creates a set of "principal experiment components" when considering the experiment as variables. This principal experiment component set indicates the experiment conditions that explain the genes behavior. In addition, when the analysis is performed on the experiment and the genes together, it results in a sub-matrix which summarizes the variance in the original data.

The singular value decomposition (SVD) (Watkins, 2004; Holter, et al., 2000) aims to uncover the underlying patters in the gene expression data. Where, the complex gene expression profiles can be captured by a small number of the characteristic modes which capture the temporal gene expression change pattern. Before applying the SVD, the rows and columns must have a zero mean. This is achieved by subtracting the mean values of the data and performing iterative normalization procedures for the rows and columns (Eisen, Spellman, Brown, & Botstein, 1998) . Given the gene expression data matrix [*A*] the SVD states that [*A*] can be written as in Eq. 1. Where the [*X*] is (*M x M*) matrix, [*Z*] is (*N x N*) matrix and [ΣZ^T] is (*M x N*) matrix and the superscript *T* means the matrix transpose. The [*X*] and [*Z*] are orthogonal. The r (the rank of matrix A is no longer than N-1) positive non-zero Eigen-values of [A^TA] are the same as Eigen-values of [AA^T]. The singular values (denoted by S_i) are the square root of these Eigen-values. The [Σ] matrix elements are all zero except for $\sum_{ii} = S_i$ for i= 1, r. The ith column Z_i of [*Z*] is the Eigen-vector of the matrix [AA^T] for i= 1, r. The ith column X_i o of (*X*) is calculated by the formula expressed in Eq. 2. Due to the form of [Σ] the other columns of [*X*] and [*Z*] matrix are irrelevant. As illustrated in Eq. 3, the jth gene temporal variation is expressed as a linear combination of the r characteristics variables. Where, the $Y_i(t)$ is the first r rows of the matrix [ΣZ^T]. It's the characteristic modes associated with matrix [A].

$$A = X\Sigma Z^T \tag{1}$$

$$X_i = \frac{1}{s_i} A Z_i \text{ For i= 1, r} \tag{2}$$

$$A_i(t) = \sum_{i=1}^{r} X_{ji} Y_i(t) \tag{3}$$

The independent component analysis (ICA) (Hyvarinen, Karhunen, & Oja, 2001) (Liebermeister, 2002) is another gene regulatory network analysis technique. The ICA is an unsupervised exploratory analysis technique which models the gene expression data matrix [*A*]using hidden variables [*Y*]as expressed in Eq. 4 and the value of each A_{il} is expressed in Eq.5.The values of [*X*] matrix are the new "independent component" variables. The statistical dependences between the columns of matrix [*X*] must be minimized. As shown in Eq.6 the statistical dependences between variables is computed by mutual information *MI*. Where, Z_k and *Z* is the entropy of the kth variable and the total entropy respectively (Cover & Thomas,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1991). Minimizing the mutual information *MI* is equivalent to minimizing the marginal entropies Z_k since the total entropy *Z* remains constant under the linear transformation. The ICA shows high fraction values around zero; therefore it can identify the approximately sparse components. To avoid this limitation Hyvarinen presented the Fast ICA algorithm (Aapo, 1999). Where, the hidden variables [*Y*] (expressed in Eq. 7) is the product of the *B* and $C^{1/2}$. The $C^{1/2}$ is the linear correlation of the data covariance matrix *C* (Walpole, Myers, Myers, & Ye, 1993). The rotation matrix *R* is initialized to random values and adjusted to minimize the statistical dependence between the independent components. The ICA suffers from the data sensitivity particularly the regulatory variables whose influence of the genes follows the Super Gaussian distribution (Liebermeister, 2002).

$$A = XY \tag{4}$$

$$A_{il} = \sum_{K} X_{ik} Y_{kl} \tag{5}$$

$$MI = \sum_{K} Z_{k} - Z \tag{6}$$

$$Y = RC^{1/2} \tag{7}$$

The gene regulatory network is driven by hidden regulatory signals. Hence, the goal of the gene regulatory network analysis technique is to uncover these hidden signals. These hidden regulatory singles are Transcription Factors (TFs) and are proteins involved in the processes of converting or transcribing DNA into RNA. The traditional computational method such as PCA and ICA ignore the underlying network structure and the decomposition is based on statistical constraints. Hence the provided model may not contain biologically or physically meaningful signals. Furthermore, these dimensional reduction techniques don't address the hidden dynamic reconstruction problems. In addition, PAC and ICA restrict the data to be mutually orthogonal and statistically independent. To avoid these limitations Liao et al. presented the Network Component Analysis (NCA) technique (Liao, et al., 2003). The mathematical model is illustrated in Eq. 4 where [*A*] is the data matrix of size (M x N) and [*X*]is the connectivity matrix of size (M x L) which encodes the connectivity strength (transcription regulation strength) between regulatory signals (TFs) and the output domain (gene expression data). In addition, the single matrix [*Y*] is of size (L x N) where L is the number of the regulatory signals which is much smaller than M. The decomposition of [*A*] into two matrices [*X*][*Y*] is an inverse problem which has no unique solution. Therefore, further assumptions are needed. These assumptions are called NCA criteria which are as follows:

- The connectivity matrix [*X*] must have a full column rank
- The connectivity matrix [*X*] must still have a full row rank when a node is removed which implies that [*X*] must have L-1 zeros
- The single matrix [*Y*] must have a full row rank

The NCA is an iterative algorithm where initial guess of the connectivity matrix [*X*] is formed by setting all the elements corresponding to missing edges between the regulatory layer and the output layer to zero moreover, the remaining elements are initialized to an arbitrary value. In addition, it suf-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

fers from the computational instability and multiple local solution problems. Therefore, Change et al. presented the FastNCA (Fast Network Component Analysis) algorithm (Chang, Ding, Hung, & Fung, 2008). FastNCA is faster than NCA since it is a non-iterative algorithm; furthermore it avoids the NCA limitations. The model of the gene regulatory network is shown in Eq. 4 and Eq.8 illustrates the model with noise. The FastNCA algorithm is illustrated in algorithm 1.

$$B = A + N = XY + N \tag{8}$$

Algorithm 1. FastNCA
Input: [A]: data matrix
Output: Gene Regulatory Network
Step 1: Perform rank-L EYM (Eckart–Young–Mirsky) approximation of *B*by SVD
Step 2: Let $W = U_L$ (left singular vector)
Step 3: Estimate connectivity matrix [*X*]
Step 3.1: For i=1 to L

Re-order rows of W such that the i[th] column of *X* has the form $a_i = \begin{bmatrix} \tilde{a_i} \\ 0 \end{bmatrix}$

Partition $W = \begin{bmatrix} W_c \\ W_r \end{bmatrix}$

Get $[V_0]$ the right singular vectors by applying SVD to W_r
Compute $\tilde{a_i}$ = the left singular vector of $W_c V_0 V_0^T$
Step 4: Estimate the [*Y*] matrix

The regulatory network is critical for understanding the pathological phenotypes and the normal cell physiology. However, the existing computational methods are successful only in the case of lower eukaryotes with simple genomes. Hence, Margolin et al. presented ARACNE (Margolin, et al., 2006) (**A**lgorithm for the **R**econstruction of **A**ccurate **C**ellular Network**s**) which is designed to scale up to the regulatory network complexity. ARACNE is Mutual Information (MI) based on gene regulatory network analysis. That eliminates the majority of the indirect interactions inferred by co-expression. ARACNE computes the MI using the Gaussian Kernel estimator (Beirlant, Dudewicz, Gyorfi, & Van der Meulen, 1997). Another MI estimator is based on B-spline function (Shi, Schmidt, Liu, & Muller-Wittig, 2011) and k-nearest neighbor distances (Kraskov, Stogbauer, & Grassberger, 2004).

Another trend of the gene regulatory network analysis technique is based on the reverse engineering concept. Where, a mathematical model is fitted to the gene expression data. The reverse engineering is the top-down approach in which the data is used to infer the parameter values for the system model. The most of recent reveres engineering techniques are based on expression data from gene expression microarrays. However, most of them don't take temporal spatial aspects of gene expression data into consideration. Nanfack et al. presented the island Evolutionary Strategy (iEs) (Fomekong-Nanfack, Kaandorp, & Blom, 2007) technique. Furthermore, some gene regulatory network analysis techniques are based on the regression algorithms (Gregoretti, Belcastro, Di Bernardo, & Oliva, 2010). Furthermore, there are some researches that focus on the biological point of view such asexploringthe functional and regulatory network of PIP co-modulated genes in human breast cancer. Marie-Anne Debily et al. (Debily,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

et al., 2009) performed gene expression profiling and network analysis studies on breast cancer cell lines which presented different PIP expression levels. The functional analysis is performed using the Ingenuity Pathway Analysis (IPA) tool. Then, a network analysis technique is performed to investigate the relationship between genes and gene products. These studies proposed that the PIP gene expression is associated with a decrease in cell proliferation, migration potential and increase of apoptotic pathways. Also, they found that the STAT5 (Signal Transducer and Activator of Transcription 5) motif has an important role in a PIP expression regulatory network.

Parallel Techniques

The gene regulatory network computational methods are computationally expensive and time consuming tasks. Since, it is affected by the huge data size. Therefore, parallel algorithms are needed to enhance the techniques performance. Gabrele Sales et al. presented a parallel package called parmigene (**PAR**allel **M**utual **I**nformation calculation for **GE**ne **NE**twork reconstruction) (Romualdi & Chiara, 2011). Parmigene is a parallel R package for mutual information and gene network reconstruction. Parmigene implement the network inference using a minimally biased MI estimator based on the k-nearest neighbor distances (Kraskov, Stogbauer, & Grassberger, 2004). Also, it uses OpenMP to implement gene regulatory network. Table 1 shows the runtime of Parmigene against the minet algorithm (Meyer, Lafitte, & Bontempi, 2008). Parmigene is 12 times faster in MI estimation and 1.3 times faster than ARACNE algorithm (Margolin, et al., 2006). The Parmigene achieved speedups of up to 50 on quad-core system.

Haixiang Shi et al. (Shi, Schmidt, Liu, & Muller-Wittig, 2011) presented accelerated B-spline functions based on mutual information (MI) estimation regulatory network algorithm on GPUs. The B-spline function based on MI estimation is accelerated using the CUDA programming model. Also, a shared memory is used for fast input/output operations moreover, a double precision floating point arithmetic and efficient partitioning schema is used to overcome the GPU device memory limitations. Haixiang Shi et al. accelerated the B-spline function based mutual information (MI) estimations by using the fine-grained model where threads in the same thread block co-operate to compute the MI value for each pair of genes. The achieved speedup is up to 82 on aNvidia Tesla C2050 GPU compared to the multi-threaded version.

Jostins et al. presented synchronous and asynchronous versions of the Parallel island Evolutionary Strategy (piES). The piES is a parallel version of the serial island Evolutionary Strategy (iEs) (Fomekong-Nanfack, Kaandorp, & Blom, 2007) developed by Fomekong et al. The island Evolutionary Strategy algorithm operates on N populations of individuals, each with a population size x. Each population is initialized independently and selection, recombination and mutation are performed only within populations. A migration operation links the population in the serial algorithm; the best individual from each island is copied to a population randomly selected from a uniform distribution. Migration occurs in every m generation.. In the parallel version each population runs on a separate processor. The speedup

Table 1. Runtime of Parmigene against the minet algorithm in minute

No. of Genes	Minet	Parmigene
100	0.02	0.005
1000	4.20	2.21
8799	2120.7	171.8

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of synchronous is 8.7 to 29.2 compared to the sequential piES. On the other hand, the speedup of the asynchronous version is 10.3 to 38.5 since the processors send and receive communication regardless the others' activity which eliminates the waiting time and leads to additional speedups. Table 2 and Figure 3 illustrate the runtime (in minutes) of the synchronous and asynchronous version of piES verses different numbers of the processing nodes.

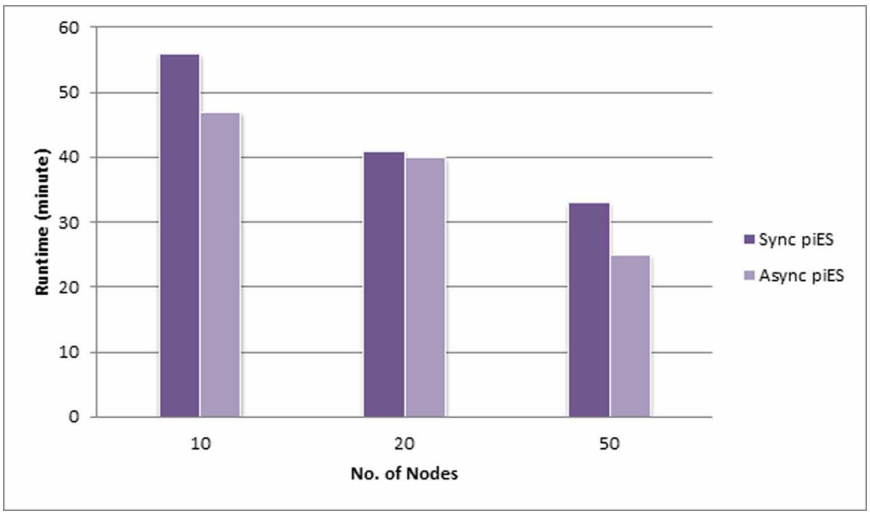
Gregoretti et al. presented a parallel network identification technique based on the Multiple Regression algorithm (NIR) (Gregoretti, Belcastro, Di Bernardo, & Oliva, 2010). The gene network is described as a system of differential equations where each equation describes the variation in time of the concentration of a particular gene as non-linear function of the concentrations of the other genes. The parallelization of this algorithm depends on assigning different genes to different computing processes equation describes the variation in time of the concentration of a particular gene as a non-linear. The speedup of algorithm is 10.1 to 102.0. Table 3 and Figure 4 show the runtimes of the NIR algorithm verses different numbers of the processing nodes.

Belcastro et al. (Belcastro, et al., 2012) presented a parallel reverse engineering gene regulatory network algorithm based on computing pair-wise Mutual Information between each gene pair. Also, communities within the gene network are discovered using a parallel hierarchical clustering. Network communities are enriched in genes involved in the same biological functions. The algorithm distributes the gene expression data across computing processors which is logically organized in a topological ring. In other words, the processing nodes are named P_0 to P_{p-1} where P_j follows P_i in case of $j > i$. In addition, P_0 follows $P_{p-1.}$ Figure 5 illustrates a ring topology example of 5 processing nodes.

Table 2. Runtime (in minutes) of the synchronous and asynchronous version of piES

No. of Nodes	Sync piES	Async piES
10	56	47
20	41	40
50	33	25

Figure 3. Runtime (in minutes) of the synchronous and asynchronous version of piES

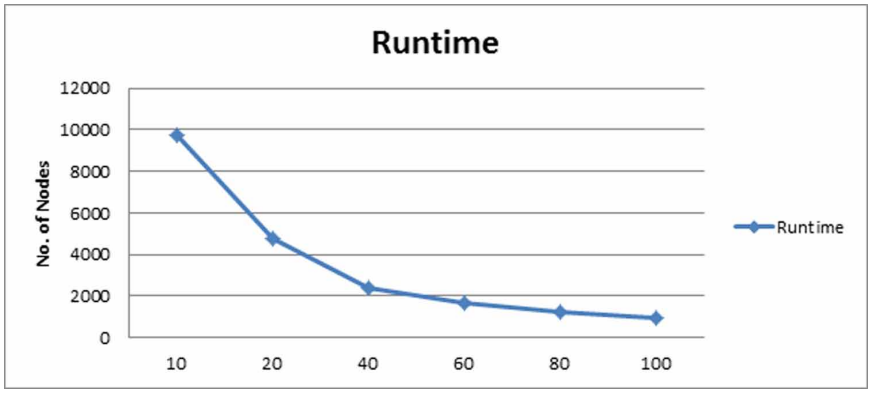


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. The NIR algorithm runtime

No. of Nodes	Runtime
10	9725
20	4798
40	2406
60	1643
80	1259
100	969

Figure 4. The NIR algorithm runtime



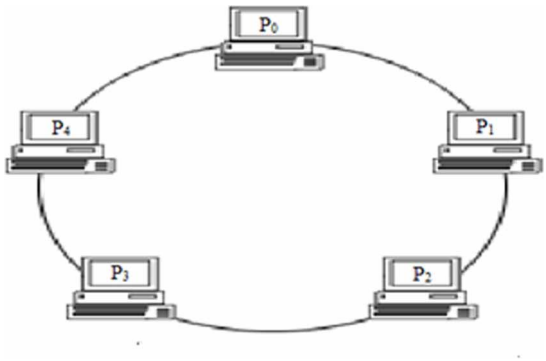
Time Series

There are researches that focus on the regulatory network construction of time series data. The construction of time series regulatory networks is the main problem in molecular biology. The time series is gene expression data measured in different time steps. It's used to identify the intermediate events between perturbation and expression responses. Also, it provides information about the chronological order of regulatory events and it provides information about the gene regulatory network edges direction. One of the time series data analyses is studying the gene response variance under different conditions.

Raychaudhuri et al. applied PCA techniques for time series data (Raychaudhuri, Stuart, & Altman, 2000) where the experiment is considered as variables. As preprocessing steps a natural log transform (Eisen, Spellman, Brown, & Botstein, 1998) is applied to avoid the influence of a gene expression level below and above one. The up-regulated genes have positive expression level while the down regulated genes have a negative expression level. Then, a covariance matrix which measures the experiments contribution in the data set is created. After that, the (n x n); where n is the number of the experiments, the covariance matrix is used to calculate the n Eigen values and their corresponding Eigenvector, where, each Eigenvector defines a principle component. The component can be considered as a weighted sum of the experiments where the weights are the Eigen-vector co-efficient. As shown in Eq. 9, the projection of the i^{th} gene is defined by the j^{th} principle component. Where the C_{kj} is the k^{th} co-efficient of the j^{th} principle component and the P_{ik} is the i^{th} gene expression level in the j^{th} experiment. In addition, the P'

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The ring topology



is the principle component data matrix, where the data is transformed from the original space to the principle component space. The larger Eigen values associated with the Eigenvector, the more contained information. On the other hand, the small Eigen values associated with the Eigenvector are uninformative. The components with less than (70/n)% of the overall variability are discarded. Applying the PCA on time series data is debatable because of the problems with the data point's dependency and the uneven time intervals.

$$P_{ij}^{'} = \sum_{k=1}^{n} P_{ik} C_{kj} \tag{9}$$

The analysis of time series data has many challenges. One being that it is hard to identify the correct sub-set of genes with a clear periodic signature. So, Wichert et al. (Wichert, Fokianos, & Strimmer, 2004) presented two statistical methods for signal detection and gene selection. Also, a statistical test was presented to identify genes relevant for the cell cycle and, a graphical tool is presented to assess whether the data contains periodic transcripts.

The genes co-expression in large data implies co-regulations and gene interactions; however it doesn't provide enough information about interaction directions. In fact, the genetic data helps the scientists to map DNA variations to RNA level variations while as mentioned before; the time series data contains information about the chronological order of regulatory events. So, Yeung et al. (Yeung, et al., 2011) presented a Bayesian model averaging regression algorithm that incorporate external information from diverse data types to infer regulatory networks from the time series and genetics data.

Jianguo et al. presented comprehensive servers for the time series metabolomic data analysis (Xia, Mandal, Sinelnikov, Broadhurst, & Wishart, 2012). The server contains many modules such as quality control, function enrichment and a metabolic path way analysis. The quality control module enables users to evaluate the data quality before conducting any analysis tasks. While the functional enrichment is useful to identify biologically meaningful patterns. Moreover, the metabolic path way analysis module is useful to analyze the pathways and visualize organisms.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

METHODOLOGY

The traditional genetics experiments that manipulate single genes or monitor the effects of a single gene in different experimental conditions are valid methods of gathering data about genes. However, they unearth very little information about other genes or how genes may interact with the collective gene product in a living organism. A useful technique that effectively transforms a living cell from black box into a transparent box is the Microarrays. They allow one to identify the genes that are expressed in different cell types, to learn how their expression levels change in different developmental stages or disease states and to identify the cellular processes in which they participate. This section illustrates the microarrays and the microarrays data analysis techniques. Besides that, the different parallelism techniques are discussed.

Microarrays

The DNA Microarrays technology is a high throughput experimental technique that can measure expression levels of hundreds of thousands of genes simultaneously. In fact, expression level of the gene is estimated by measuring the amount of mRNA for that gene. Actually, the procedure of a DNA microarray experiment includes multiple steps from sample preparation to data analysis (Figure 6), among which hybridization is a central process. Where, the sample that contains the targets to be investigated is added to the DNA chip to allow their binding to the probes fixed on the chip, resulting in a characteristic binding pattern representing the levels of gene expression of the sample. The sample itself is labeled prior to the hybridization and the mostoftenused labels are fluorescence labels that allow detection of the binding event. After the hybridization, the chip is washed and then fluorescence intensities on the chip are read and recorded by a scanning or imaging device. Raw data reflecting the fluorescence intensities are statistically analyzed and often shown by fold changes as compared to control (Li & Li, 2008).

The first step in microarrays experiment is the sample preparation where a microarray is prepared either specifically for the genes of interest or for all of the genes in the organism of interest's genome. This is done by synthesizing a complementary probe to the mRNA for each gene. Namely, the most common technique for doing this is to use robotic systems to individually synthesize thousands of probes, with one spot containing thousands of oligonucleotide probes for a single gene. After that, you can extract the mRNA from cells you wish to analyze, this mRNA represents all of the expressed genes in those cells and then the mRNA is labeled with a fluorescent dye. The second step is Hybridization where both the experimental and the control mRNA are simultaneously washed over the microarray. Whenever the probes on a given spot interact with a complementary mRNA strand, the two will hybridize, thus binding the mRNA to the microarray. After a certain period of time the remaining solution is discarded and the microarray is gently rinsed. In the third step; Dictation, the microarray is scanned optically using Microarrays Scanner and the ratios of red to green are computed. Since red represents the control mRNA and green represents the experimental mRNA, the colors on each spot will reflect how the expression of each gene was changed by the manipulation of the experimental condition. Often a new image of the microarray is formed almost exactly the same as the original, with the exception that now green means a reduction in expression and red means an increase in expression. The final step is the data acquisition and analyses of the red: green ratios which are analyzed mathematically for relationships among the genes (Cheung, et al., 1999).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The analysis of the data resulting from microarrays remains a big challenge for the huge volume of data it produces. In fact, the data analysis process involves various computational tasks (Yang, et al., 2008). As shown in Figure 7 the microarrays data analysis process starts with a sub-set extraction. In this step, the dataset is filtered to extract the differentially expressed genes, (discriminator genes) which are the genes with significantly different expression into two defined groups of microarray experiments. The output of this step can be the input of other analysis tasks such as component analysis, clustering and network analysis. The aim of a gene expression search is to find a set of genes that have similar expression patterns with a target gene. Actually, when the function of a target gene is unknown, inspecting the function of genes with expression patterns similar to the target gene may reveal the function of the target gene. On the other hand, when the function of the target gene is known, we can have an insight into how similarly/dissimilarly genes are expressed in given conditions.

The microarrays data clustering aims to organize genes where those with similar expression patterns are grouped together to identifying biologically relevant groups of genes. While, the aim of component analysis is studying the ways in which genes responses varies under different conditions. Moreover, network analysis (Haman & Valenta, 2013) is a powerful analysis tool to understand real world phenomena. There are two types of regularity network, gene regularity network and PPI (protein-protein interaction) network. The gene regulatory network or genetic regulatory network (GRN) is a group of DNA segments which interact with each other through their RNA and protein expression products in an indirect way.

Figure 6. Workflow of a DNA microarray experiment

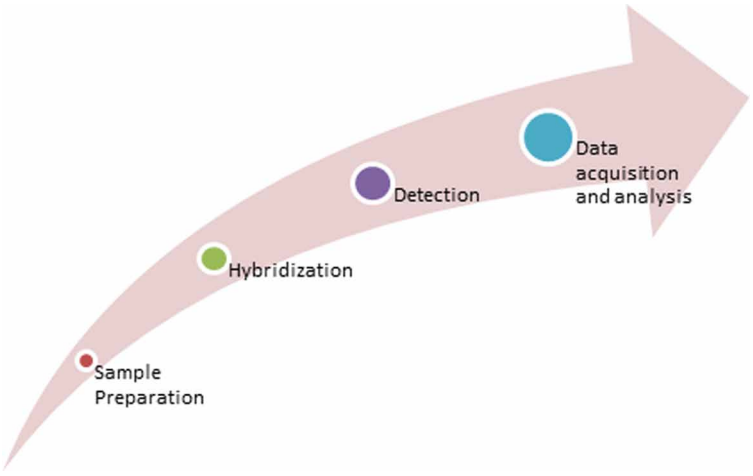
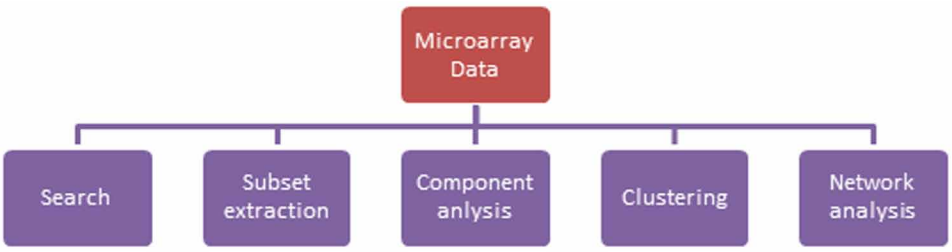


Figure 7. Microarrays data analysis tasks



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Also, they interact with the other substances in the cell, actually, these interactions governing the gene transcription rates into mRNA. On the other hand, the protein-protein interaction network plays a key role in understanding the protein docking process and the cellular system working method, due to the huge size of data sets produced by microarray technology, analysis tasks are usually computationally intensive and time-consuming. Therefore, it requires high-performance computing power.

Parallelism Techniques

In the past decades it was believed that program execution time was improved by creating faster and more efficient processors to deal with the growing massive volumes of the data. However, parallel computing was introduced to execute programs in parallel instead of sequential executions (Kumar, Grama, Gupta, & Karypis, 2003). The program is divided into tasks and the independent tasks are executed at the same time on different processing nodes. In this way, the parallel computing concept succeeded to minimize the cost of building expensive hardware since it benefits from the computing power of different processors instead of increasing the computing power of one processor. Parallel computing models are categorized into two areas: Process interaction and problem decomposition. The process interaction focuses on how the processes communicate with each other. While, the problem decomposition focuses on how the program is divided into independent tasks.

The problem decomposition is classified into two trends: Task parallelism and data parallelism. On one hand, the task parallelism focuses on the decomposition of the program into independent tasks which are executed simultaneously. While, the data parallelism focuses on performing operations on a dataset. Where, a set of tasks will operate on separate partitions of the data independently. In the shared memory model, the data will be shared between the processes while in the distributed memory the data will be divided between the memories of the processes.

The process interaction is categorized into two main categories; shared memory architecture (Karp, 1988) (Hager & Wellein, 2010) and the distributed memory architecture (Hager & Wellein, 2010) (also called message passing architecture). As shown in Figure 8-a, in the distributed memory architecture the memory is distributed and the processes communicate by sending messages. These communications can be asynchronous or synchronous. On the other hand in the shared memory architecture, the memory is shared among all the processes and the processes communicate implicitly by updating shared variables (Figure 8-b). The shared memory model requires protection mechanisms such as locks, semaphores and monitors to control concurrent access. Sometimes, the two architectures can be merged together to benefit from the full computing power of the hardware, this architecture is called hybrid architecture as shown in Figure 8-c. In this case, the computational tasks are distributed among the processing nodes and each node executes its own task on a number of threads.

Another form of the distributed memory architecture is cloud computing (Erl, Puttini, & Mahmood, 2013) where groups of remote servers are deployed (Figure 8-d). The main advantage of cloud computing is that the user focuses only on the business core instead of being involved in the hardware management. The deployment model of cloud computing is classified into three categories private, public and hybrid. In the private cloud the infrastructure is operated for a single organization. On the other hand, in the public cloud the infrastructure is rendered for public usage, while, the hybrid model is a collection of the private and the public cloud models.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The graphic processing unit (GPU) (also called visual processing unit (VPU)) can be used to enhance the performance. The GPU is an electronic circuit used to accelerate image creation and manipulation for the output unit's and it's more efficient than the CPU (central processing unit) because of its parallel architecture. The famous GPU programming language is CUDA (Compute United Device Architecture) (Sanders & Kandrot, 2010) and OpenCL (Open Computing Language) (Gaster, Howes, Kaeli, Mistry, & Schaa, 2012). As shown in Figure 9 the CUDA execution model consists of a grid where the grid is a set of blocks and each block is a set of concurrent threads and each thread has a unique ID. The threads in the same block communicate through the per block shared memory, in addition, the threads in the same block are synchronized using barriers. The threads in different blocks have no direct communication and cannot be synchronized directly. CUDA supports two types of parallelism, the coarse-grained and fine-grained. In the coarse-grained, a large number of threads run independently in parallels, whereas in the fine-grained, the threads located in the same thread block co-operate simultaneously on some shared memory space. As illustrated in Figure 10 the OpenCL execution model consists of Index space which consists of a set of work groups. Each workgroup contains a number of independent work items.

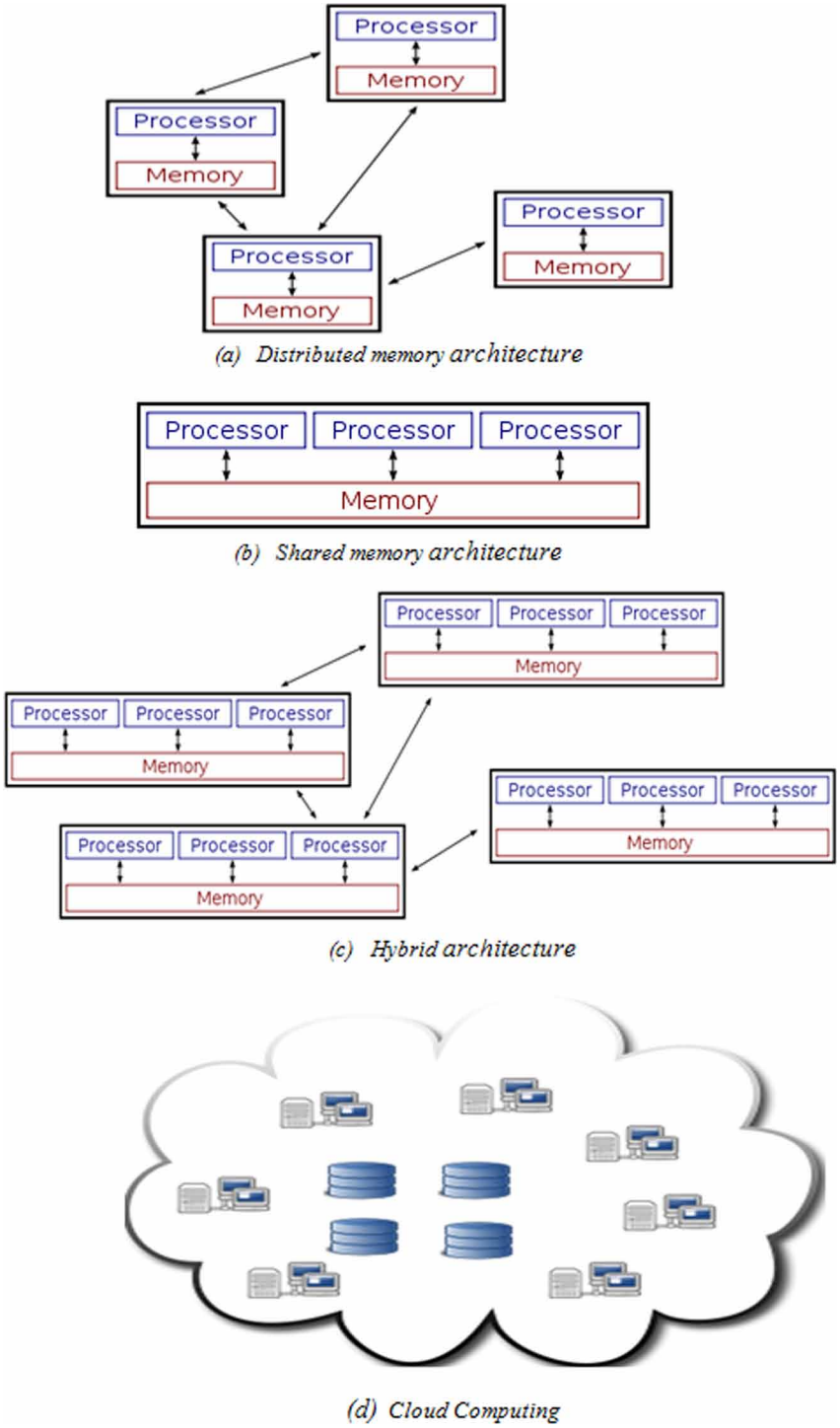
CHALLENGES

One of the microarrays data analysis tasks is the gene regulatory network. The gene regulatory network aims to study the genes interactions topological organization and gene-to-gene influences. Where each gene is represented as a node and regulatory interactions between genes is represented as edges. The gene regulatory can be classified into four categories; component analyses, mutual information, reverse engineering and regression techniques. One of the component analysis techniques is the principle component analysis (PCA) technique (Raychaudhuri, Stuart, & Altman, 2000). The PCAis a dimensionality reduction technique which determines the key variables that explain the difference in the data and simplify the data visualization and analysis. The PCA assumes that the data is orthogonal which is not always valid in biological data. Another trend is the singular value decomposition (SVD) technique (Watkins, 2004) (Holter, et al., 2000)which aims to uncover the underlying patters in the gene expression data. Where, the complex gene expression profiles can be captured by a small number of the characteristic modes which capture the temporal gene expression change pattern. SVD suffers from the noise sensitivity also some results aren't meaningful. In addition, another trend is the independent component analysis (ICA) which may provide multiple local solutions. To avoid the limitations of PCA, SVD and ICA Liao et al. presented the Network Component Analysis (NCA) technique (Liao, et al., 2003).Although that, it is computational expensive and provides multi-local solutions. In addition, Change et al. presented the FastNCA (Fast Network Component Analysis) algorithm (Chang, Ding, Hung, & Fung, 2008) which is faster than NCA. However, the FastNCA is a sequential algorithm.

The regulatory network is critical for understanding the pathological phenotypes and the normal cell physiology. Mutual information (MI) based techniques is another trend for the gene regulatory network analysis.Margolin et al. presented ARACNE (Margolin, et al., 2006) (**A**lgorithm for the **R**econstruction of **A**ccurate **C**ellular Network**s**) which is designed to scale up to the regulatory network complexity. ARACNE computes the MI using the Gaussian Kernel estimator (Beirlant, Dudewicz, Gyorfi, & Van der Meulen, 1997). Another MI estimator is based on B-spline function (Shi, Schmidt, Liu, & Muller-Wittig, 2011) and k-nearest neighbor distances (Kraskov, Stogbauer, & Grassberger, 2004). The drawback of the MI based techniques is biased to the mutual information estimation method.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Parallel computing architecture



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. The CUDA execution model

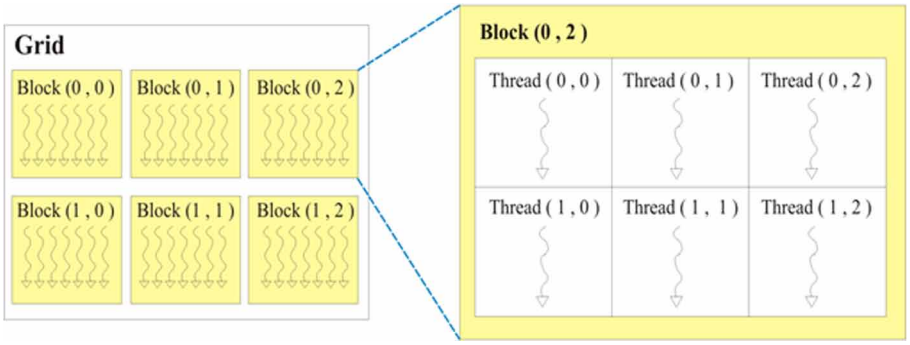
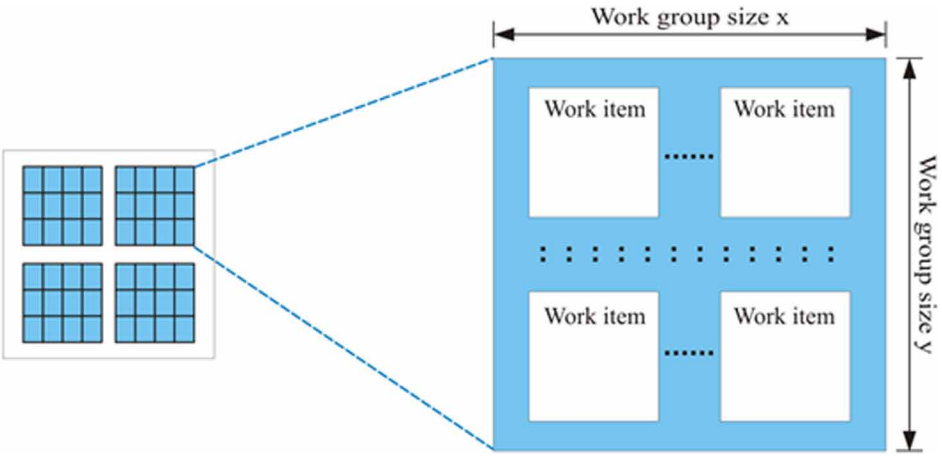


Figure 10. The OpenCL execution model



Another trend of the gene regulatory network analysis technique is based on the reverse engineering concept. Where, a mathematical model is fitted to the gene expression data. The reverse engineering is the top-down approach in which the data is used to infer the parameter values for the system model. The reverse engineering techniques assume that the data is linear which is always valid in the biological data. In addition, In addition, these techniques don't take the temporal spatial aspects of gene expression data into consideration.

The regression based techniques is another trend of the gene regulatory network analysis techniques. Where, the gene network is described as a system of differential equations where each equation describes the variation in time of the concentration of a particular gene as non-linear function of the concentrations of the other genes. The drawback of these techniques is the noise sensitive. Furthermore, the commons drawback of these techniques is the computational expensive where the techniques runtime is affected by the size of data. The drawbacks of the gene regulatory analysis techniques are summarized in Table 4. Since, almost all of these techniques are computationally intensive and time-consuming, parallel algorithms are needed to enhance the techniques performance.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Gene regulatory analysis techniques drawbacks

Technique	Data Assumptions	Non-Meaningful Results	Local Solutions	Data Sensitivity	Computational Expensive
PCA	✓				✓
SVD		✓		✓	✓
ICA			✓		✓
NCA			✓		✓
MI				✓	✓
Reverse Eng.	✓				✓
Regression				✓	✓

PROPOSED ARCHITECTURE

The gene expression data analysis plays important role in understanding the gene functions and relationships. One of the gene expression data analysis tasks is the gene regulatory network which aims to study how the genes interact with each other and the gene-to-gene influences. Moreover, the regulatory network is critical for understanding the pathological phenotypes and the normal cell physiology. Unfortunately, most of the recent techniques are effected be the huge data size. Therefore, this work proposes solution for gene regulatory network using a high performance technique to enhance both the speed and the efficiency called PAGeneRN (**P**arallel **A**rchitecture for **GeneR**egulatory **N**etwork).In fact, the parallel techniques can be divided into two categories based on the parallelism type. Some algorithms are functionality paralyzed where each processing node executes different task simultaneously. On the other hand, other algorithms are data paralyzed where each processing node execute the same function on a chunk of the data simultaneously.

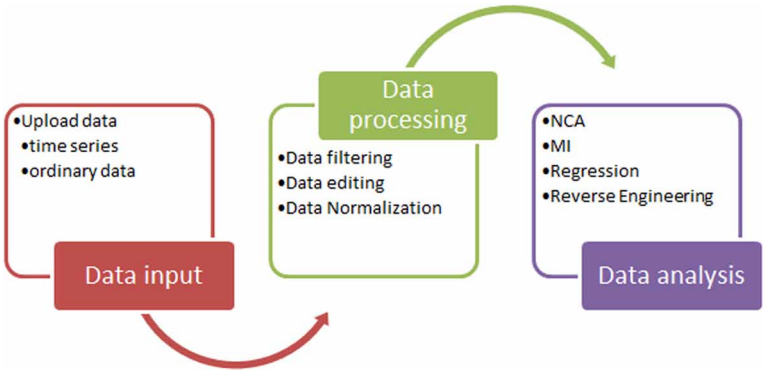
The proposed architecture (PAGeneRN) will combine the two different parallelism methodologies (data and functionality parallelism model) to take the full advantage of the high performance computing power. In addition, the proposed algorithm will use the computational power of the GPUs. Furthermore, a complete gene regulatory analysis framework is proposed. As shown in Figure 11 the proposed framework consists of three main modules data input, data processing and finally the data analysis module.

The Data input module enable user to input the data to the system. The data may be ordinary gene expression data or time-series. The next module is the data processing which aims to transform the raw data into cleaner data through set of procedures such as filtering and data editing. In other words, the data processing module has different options such as data filtering, data editing and data normalization option. Data filtering option enables user to remove empty or low quality data points from the dataset. On the other hand, the data editing option enable user to delete/edit data points. While, the data normalization is important for the biological data analysis where some techniques assume that the data is normally distributed. Also, data normalization reduces the systematic bias in the data. The last module is the data analysis module which is the core module. The data analysis module consists of different parallel gene regulatory network analysis techniques.

The data analysis module consists of a parallel version of the FastNCA algorithm (Chang, Ding, Hung, & Fung, 2008) (illustrated in algorithm 2 and 3). Where, the master processing node will distribute the data among the worker node. Where each worker will take a chunk of the data of size N/P (N is data

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Proposed architecture modules



size). In addition, the framework will contain Mutual information based technique where the mutual information estimated using the cubic B-spline. In addition, the framework will contain parallel version of regression and reverse engineering techniques. Moreover, these techniques can be combined with the genetic information to avoid the meaningless results problem.

Algorithm 2. FastNCA parallel version – master node
Input: [A]: data matrix
Output: Gene Regulatory Network
Step 1: Perform rank-L EYM (Eckart–Young–Mirsky) approximation of *B* by SVD
Step 2: Let W = U_L (left singular vector)
Step 3: Estimate connectivity matrix [*X*] in parallel
Step 3.1: Send chunk of the data for each processing node
Step 3.2: Collect data from the processing nodes
Step 4: Estimate the [*Y*] matrix
Algorithm 3. FastNCA parallel version – worker node
Input: [W]: left singular vector matrix
Output: Connectivity sub matrix [*X*]
Step 1: Parallel for i=1 to L

Re-order rows of W such that the i[th] column of *X* has the form $a_i = \begin{bmatrix} \tilde{a_i} \\ 0 \end{bmatrix}$

Partition $W = \begin{bmatrix} W_c \\ W_r \end{bmatrix}$

Get [V_0] the right singular vectors by applying SVD to W_r
Compute $\tilde{a_i}$ = the left singular vector of $W_c V_0 V_0^T$
Step 2: Send connectivity sub matrix [*X*] to the master node

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

The gene regulatory network is one of the microarrays data analysis tasks which aims to study the gene-to-gene influences and genes interactions topological organization. Where each gene is represented as a node and regulatory interactions between genes is represented as edges. Many researches focused on the gene regulatory network analysis. This work presents a background, mathematical models and comparisons about the gene regulatory networks analysis different techniques. The gene regulatory is classified into four categories; component analyses, mutual information, reverse engineering and regression techniques. Furthermore, the component analysis techniques are categorized as principle component analysis (PCA), singular value decomposition (SVD), independent component analysis (ICA) and network component analysis (NCA) technique. The PCA and Mutual information based techniques suffers from assuming that the data is orthogonal and linear respectively. On the other hand, the SVD technique sometimes provides non-meaningful result also it suffers from the noise sensitivity. Furthermore, the ICA may provide multiple local solutions. However, the NCA avoids the problems of the PCA, SVD and ICA techniques, it is a computational expensive. In addition the mutual information based techniques is biased to the mutual information estimation method. Furthermore, the regression technique is noise sensitive. This paper also proposes PAGeneRN (**P**arallel **A**rchitecture for **Gene R**egulatory **N**etwork) gene regulatory analysis architecture which uses a parallel model to enhance the technique performance. The proposed architecture consists of three main modules: Data input, data processing and the data analysis module. The Data input module enable user to input the data to the system. While, the data processing aims to transform the raw data into cleaner data through set of procedures such as filtering and data editing. Furthermore, the core module is data analysis module which consists of different parallel gene regulatory network analysis techniques.

REFERENCES

Aapo, H. (1999). Fast and robust fixed-point algorithms for independent component analysi. *IEEE Transactions on Neural Networks*, *10*(3), 626–634. doi:10.1109/72.761722 PMID:18252563

Aluru, S. (2006). *Handbook of computational molecular biology*. CRC Press.

Beirlant, J., Dudewicz, E. J., Gyorfi, L., & Van der Meulen, E. C. (1997). Nonparametric entropy estimation: An overview. *International Journal of Mathematical and Statistical Sciences*, *6*(1), 17–39.

Belcastro, V., Gregoretti, F., Siciliano, V., Santoro, M., DAngelo, G., Oliva, G., & di Bernardo, D. (2012). Reverse engineering and analysis of genome-wide gene regulatory networks from gene expression profiles using high-performance computing. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *9*(3), 668–678. doi:10.1109/TCBB.2011.60 PMID:21464509

Berrozpe, G., Schaeffer, J., Peinado, M., Real, F., & Perucho, M. (1994). Comparative analysis of mutations in the p53 and K-ras genes in pancreatic cancer. *International Journal of Cancer*, *58*(2), 185–191. doi:10.1002/ijc.2910580207 PMID:8026879

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chang, C., Ding, Z., Hung, Y. S., & Fung, P. C. (2008). Fast network component analysis (FastNCA) for gene regulatory network reconstruction from microarray data. *Bioinformatics (Oxford, England)*, *24*(11), 1349–1358. doi:10.1093/bioinformatics/btn131 PMID:18400771

Cheung, V., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R., & Childs, G. (1999). Making and reading microarrays. *Nature Genetics*, *21*(1), 15–19. doi:10.1038/4439 PMID:9915495

Churchill, G. (2002). Fundamentals of experimental design for cDNA microarrays. *Nature Genetics, 32*(1), 490 - 495.

Cosmides, L., & Tooby, J. (1995). *From function to structure: The role of evolutionary biology and computational theories in cognitive neuroscience*. The MIT Press.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. John Wiley & Sons. doi:10.1002/0471200611

Culf, A., Cuperlovic-Culf, M., & Ouellette, R. (2006). Carbohydrate microarrays: Survey of fabrication techniques. *OMICS: A Journal of Integrative Biology*, *10*(3), 289–310. doi:10.1089/omi.2006.10.289 PMID:17069509

Debily, M.-A., El Marhomy, S., Boulanger, V., Eveno, E., Mariage-Samson, R., Camarca, A., ... Imbeaud, S. (2009). A functional and regulatory network associated with PIP expression in human breast cancer. *PLoS ONE*, *4*(3), e469–e486. doi:10.1371/journal.pone.0004696 PMID:19262752

Dominguez, C., Boelens, R., & Bonvin, A. M. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–1737. doi:10.1021/ja026939x PMID:12580598

Dowsey, A. W. (2005). *High-throughput image analysis for proteomics*. Citeseer.

Durbin, R. (1998). *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge University Press. doi:10.1017/CBO9780511790492

Eisen, M., Spellman, P., Brown, P., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 14863–14868. doi:10.1073/pnas.95.25.14863 PMID:9843981

Erl, T., Puttini, R., & Mahmood, Z. (2013). *Cloud Computing: Concepts, Technology & Architecture*. Pearson Education.

Fomekong-Nanfack, Y., Kaandorp, J. A., & Blom, J. (2007). Efficient parameter estimation for spatio-temporal models of pattern formation: Case study of Drosophila melanogaster. *Bioinformatics (Oxford, England)*, *23*(33), 56–63. PMID:17893088

Gasch, A., Spellman, P., Kao, C., Carmel-Harel, O., Eisen, M., Storz, G., ... Brown, P. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell*, *11*(12), 4241–4257. doi:10.1091/mbc.11.12.4241 PMID:11102521

Gaster, B., Howes, L., Kaeli, D. R., Mistry, P., & Schaa, D. (2012). Heterogeneous Computing with OpenCL [Newnes.]. *Revised Open*, *CL*, 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., ... Weissman, J. S. (2003). Global analysis of protein expression in yeast. *Nature*, *425*(6959), 737–741. doi:10.1038/nature02046 PMID:14562106

Gregoretti, F., Belcastro, V., Di Bernardo, D., & Oliva, G. (2010). A parallel implementation of the network identification by multiple regression (NIR) algorithm to reverse-engineer regulatory gene networks. *PLoS ONE*, *5*(4), e10179–e10183. doi:10.1371/journal.pone.0010179 PMID:20422008

Haefner, J. W. (2005). *Modeling Biological Systems: Principles and Applications*. Springer Science.

Hager, G., & Wellein, G. (2010). *Introduction to high performance computing for scientists and engineers*. CRC Press. doi:10.1201/EBK1439811924

Haman, J., & Valenta, Z. (2013). Shrinkage Approach for Gene Expression Data Analysis. *EJBI*, *9*(3), 2–8.

Holter, N. S., Mitra, M., Maritan, A., Cieplak, M., Banavar, J. R., & Fedoroff, N. V. (2000). Fundamental patterns underlying gene expression profiles: Simplicity from complexity. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(15), 8409–8414. doi:10.1073/pnas.150242097 PMID:10890920

Hyvarinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. John Wiley & Sons. doi:10.1002/0471221317

Janssen, P. J., Jones, W. A., Jones, D. T., & Woods, D. R. (1988). Molecular analysis and regulation of the glnA gene of the gram-positive anaerobe Clostridium acetobutylicum. *Journal of Bacteriology*, *170*(1), 400–408. doi:10.1128/jb.170.1.400-408.1988 PMID:2891680

Jostins, L., & Jaeger, J. (2010). Reverse engineering a gene network using an asynchronous parallel evolution strategy. *BMC Systems Biology*, *4*(1), 17–33. doi:10.1186/1752-0509-4-17 PMID:20196855

Karp, R. (1988). *A survey of parallel algorithms for shared-memory machines*. University of California.

Kelley, L., MacCallum, R., & Sternberg, M. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *Journal of Molecular Biology*, *299*(2), 501–522. doi:10.1006/jmbi.2000.3741 PMID:10860755

Kononen, J., Bubendorf, L., Kallionimeni, A., Börlund, M., Schraml, P., Leighton, S., ... Kallionimeni, O. (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nature Medicine*, *4*(7), 844–847. doi:10.1038/nm0798-844 PMID:9662379

Kraskov, A., Stogbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *69*(6), 66138–66154. doi:10.1103/PhysRevE.69.066138 PMID:15244698

Kumar, V., Grama, A., Gupta, A., & Karypis, G. (2003). *Introduction to parallel computing*. Addison Wesley.

Li, S., & Li, D. (2008). *DNA microarray technology and data analysis in cancer research*. World Scientific Pub Co Inc.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liao, J. C., Boscolo, R., Yang, Y.-L., Tran, L. M., Sabatti, C., & Roychowdhury, V. P. (2003). Network component analysis: Reconstruction of regulatory signals in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *20*(26), 15522–15527. doi:10.1073/pnas.2136632100 PMID:14673099

Liebermeister, W. (2002). Linear modes of gene expression determined by independent component analysis. *Bioinformatics (Oxford, England)*, *18*(1), 51–60. doi:10.1093/bioinformatics/18.1.51 PMID:11836211

Ma, H., & Horiuchi, K. (2006). Chemical microarray: A new tool for drug screening and discovery. *Drug Discovery Today*, *1*(13), 661–668. doi:10.1016/j.drudis.2006.05.002 PMID:16793536

Margolin, A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R., & Califano, A. (2006). ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*, *7*(Suppl 1), S7–S22. doi:10.1186/1471-2105-7-S1-S7 PMID:16723010

Meenakshisundaram, K., Carmen, L., Michela, B., Diego, D., Rosaria, V., & Gabriella, M. (2009). Existence of snoRNA, microRNA, piRNA characteristics in a novel non-coding RNA: x-ncRNA and its biological implication in Homo sapiens. *Journal of Bioinformatics and Sequence Analysis*, *1*(2), 31–40.

Meyer, P. E., Lafitte, F., & Bontempi, G. (2008). minet: AR/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinformatics*, *9*(1), 461–471. doi:10.1186/1471-2105-9-461 PMID:18959772

Nair, A. (2007). Computational Biology & Bioinformatics - A gentle Overview. *Communications of Computer Society of India*, *30*(1), 7–12.

Raychaudhuri, S., Stuart, J. M., & Altman, R. B. (2000). Principal components analysis to summarize microarray experiments: application to sporulation time series. In *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* (pp. 455-466). NIH Public Access.

Rivas, L., García-Villadangos, M., Moreno-Paz, M., Cruz-Gil, P., Gómez-Elvira, J., & Parro, V. (2008). A 200-Antibody Microarray Biochip for Environmental Monitoring: Searching for Universal Microbial Biomarkers through Immunoprofiling. *Analytical Chemistry*, *80*(21), 7970–7979. doi:10.1021/ac8008093 PMID:18837515

Romualdi, G. S., & Chiara. (2011). Parmigene—A parallel R package for mutual information estimation and gene network reconstruction. *Bioinformatics*, *27*(13), 1876-1877.

Rubin, G., Yandell, M., Wortman, J., Gabor, G., Nelson, C., Hariharan, I., ... Fleischmann, W. (2000). Comparative genomics of the eukaryotes. *Science*, *287*(5461), 2204–2215. doi:10.1126cience.287.5461.2204 PMID:10731134

Sanders, J., & Kandrot, E. (2010). *CUDA by example: an introduction to general-purpose GPU programming*. Addison-Wesley Professiona.

Schena, M., Shalon, D., Davis, R., & Brown, P. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, *270*(5235), 467–470. doi:10.1126cience.270.5235.467 PMID:7569999

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shi, H., Schmidt, B., Liu, W., & Muller-Wittig, W. (2011). Parallel mutual information estimation for inferring gene regulatory networks on GPUs. *BMC Research Notes*, *4*(1), 189–194. doi:10.1186/1756-0500-4-189 PMID:21672264

Shortle, D. (2000). Prediction of protein structure. *Current Biology*, *10*(2), 49–51. doi:10.1016/S0960-9822(00)00290-6 PMID:10662677

Stoevesandt, O., Taussig, M., & He, M. (2009). Protein microarrays: High-throughput tools for proteomics. *Expert Review of Proteomics*, *6*(2), 145–157. doi:10.1586/epr.09.2 PMID:19385942

Velculescu, V. E., Zhang, L., Vogelstein, B., & Kinzler, K. W. (1995). Serial analysis of gene expression. *Science*, *270*(5235), 484–487. doi:10.1126cience.270.5235.484 PMID:7570003

Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (1993). *Probability and statistics for engineers and scientists*. Macmillan New York.

Watkins, D. (2004). *Fundamentals of matrix computations* (Vol. 64). John Wiley & Sons.

Wichert, S., Fokianos, K., & Strimmer, K. (2004). Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics (Oxford, England)*, *20*(1), 5–20. doi:10.1093/bioinformatics/btg364 PMID:14693803

Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D., & Wishart, D. S. (2012). MetaboAnalyst 2.0 — a comprehensive server for metabolomic data analysis. *Nucleic Acids Research*, *40*(W1), W127–W133. doi:10.1093/nar/gks374 PMID:22553367

Yang, Y., Choi, J., Choi, K., Pierce, M., Gannon, D., & Kim, S. (2008). BioVLAB-Microarray: Microarray Data Analysis in Virtual Environment. *IEEE Fourth International Conference on eScience.* 10.1109/eScience.2008.57

Yeung, K. Y., Dombek, K. M., Lo, K., Mittler, J. E., Zhu, J., Schadt, E. E., ... Raftery, A. E. (2011). Construction of regulatory networks using expression time-series data of a genotyped population. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(48), 19436–19441. doi:10.1073/pnas.1116442108 PMID:22084118

Zhou, Q., Chipperfield, H., Melton, D., & Wong, W. (2007). A gene regulatory network in mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(42), 16438–16443. doi:10.1073/pnas.0701014104 PMID:17940043

ENDNOTE

[1] Image Source: http://www.pnas.org/content/104/42/16438/F3.expansion.html

This research was previously published in the Handbook of Research on Machine Learning Innovations and Trends edited by Aboul Ella Hassanien and Tarek Gaber; pages 762-786; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 56
Research Perspectives on Functional Micro and Nano Scale Coatings:
New Advances in Nanocomposite Coatings for Severe Applications

Jaime Andrés Pérez Taborda
Universidad Tecnológica de Pereira, Colombia & Instituto de Microelectrónica de Madrid-CSIC, Colombia

Elvis O. López
Centro Brasileiro de Pesquisas Físicas, Brazil

ABSTRACT

Research topics related to the production of nanocomposites are the most important directions of development of new semiconductor engineering, ensuring high nanocomposites obtaining useful properties in the scope of biophysical characteristics, biomedical and piezoelectric applications. We present two case studies as Hydroxyapatite are in medical applications and aluminum nitride as acoustic wave sensor. Hydroxyapatite, is the main inorganic structure of the tooth enamel and bone and is a biomaterial that is commonly used in biomedical applications that involve bone substitution, drug delivery and bone regeneration because of its excellent biocompatibility, high bioactivity and good osseoconductivity. Since the past decade. Aluminum nitride (AlN), an electrical insulating ceramic with a wide band gap of 6.3 eV, is a potentially useful dielectric material very important in fields such as optoelectronic and micro electronics.

DOI: 10.4018/978-1-7998-1204-3.ch056

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

In recent years, there has been a breakthrough in the research of new materials with increasingly complex structures with properties of ionic and electronic conduction (Morkoc, 2013). Such materials have applications in important areas of technology as components of many optoelectronic devices - among the new materials studied highlights the metal nitrides - and hydroxyapatite nanostructured films fabricated by RF magnetron sputtering for biomedical applications (Zhang, 2013).

In general, semiconductor nitrides are the group of inorganic compounds consisting of nitrogen and other more electropositive element, except for halogens and hydrogen. Nitrides of elements such as boron, titanium, zirconium, aluminum, and silicon are hard, stable at high temperatures, and are resistant to chemical agents. At a much smaller scale, coatings are used in numerous electronic products (both consumer and industrial electronics) and biomedical products.

Currently, our society is increasingly immersed in nanotechnology. They are shown, two approaches to nanocoatings applications that impact our daily lives. In section 1, a new method of magnetic confinement of plasmas RF hydroxyapatite has shown to produce a high crystalline quality for biomedical applications. In section 2, aluminum nitride nanostructures grown via pulsed laser deposition - PLD that has been used as substrates for the manufacture of acoustic wave devices - SAW are highly demanded by the telecommunications industry.

1. HYDROXYAPATITE NANOSTRUCTURED FILMS FABRICATED BY RF MAGNETRON SPUTTERING FOR BIOMEDICAL APPLICATIONS

Hydroxyapatite - HA ($Ca_{10}(PO_4)_6(OH)_2$),(Morgan, Wilson, Elliott, Dowker, & Anderson, 2000) is the main inorganic structure of the bone and enamel teeth and also is the main bioceramic that is commonly used in biomedical applications that involve drug delivery, (Hollister, 2005; Vallet-Regí & Ruiz-Hernández, 2011) bone repair and bone regeneration due its excellent properties of osteoinduction, osteoconduction and osseointegration (Astala & Stott, 2008; Paul & Sharma, 2006). Since past decades, the HA has been used as a bioactive coatings to improve the osteoconductivity of metallic implants for the orthopaedic applications. Metals as titanium and alloys have not the bioactive capacity, different deposition techniques to produce calcium phosphate (CaP) coatings has been developed to improve the osseointegration, osteoconduction and osteoinduction of the metallic implants (Dos Santos et al., 2010; Macaskie et al., 2005; Socol et al., 2004; Ueda, Narushima, Goto, Taira, & Katsube, 2007; Wang et al., 2006; Wen, Xu, Hu, & Hodgson, 2007).

Chemical and physical deposition processes are used to synthesize HA as coatings over a metallic surface, these deposition techniques are: biomimetics, (Uchida, Oyane, Kim, Kokubo, & Ito, 2004) sol-gel routes, (Melde & Stein, 2002; Padilla, Roman, Carenas, & Vallet-Regı, 2005; K. Zhang, Francis, Yan, & Stein, 2005) electrochemical deposition,(Liu, Savino, & Yates, 2011; Narayanan, Seshadri, Kwon, & Kim, 2008; Wang, Eliaz, & Hobbs, 2011) and others,(Li, Hong, Yu, & Qi, 2010; Li & Qi, 2008) as well as plasma spray, (Dong, Khor, Quek, White, & Cheang, 2003; Ellies, Nelson, & Featherstone, 1992; Khor, Dong, Quek, & Cheang, 2000) Pulsed Laser Deposition (PLD), (Carradò, 2011; Joanni et al., 2004) Ion and Electron Beam Deposition (IBD, EBD),(Jeong, Choe, & Eun, 2011; Rabiei et al., 2006) Electron Cyclotron Resonance deposition (ECR)(Akazawa & Ueno, 2013, 2014) and Radio Frequency Magnetron Sputtering (RF-MS) (Akazawa & Ueno, 2014; Boyd, Meenan, & Leyland, 2006). The chal-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

lenge is produce a thin, dense, adherent, stoichiometric and crystalline HA coating over the implant surface, thus the HA coatings will be stable in a biological environments,(Bauer, Schmuki, von der Mark, & Park, 2013) and not to degrade along the time when inside of the human body. Besides, all deposition techniques has been adapted to produce thin coatings with thickness less than of 500 nm to improve the adhesion to the substrate and to preserve the original topography of the metallic implants (Akazawa & Ueno, 2013). Nevertheless, it is difficult produce a thin HA coating without the formation of undesirables amorphous calcium phosphate phases (ACP) (Ferro, Barinov, Rau, Teghil, & Latini, 2005; Hahn et al., 2011; Surmenev, 2012; Yonggang, Wolke, Yubao, & Jansen, 2007; Zhang, 2013). Usually, CaP coatings are commonly treated at temperatures up to 900º C to transform the amorphous phases (ACP) in a HA structure (Akazawa & Ueno, 2013; Surmenev, 2012; S Zhang et al., 2007). In addition, these thermal treatment may produce undesirables phases mainly as tricalcium phosphate (α,β –TCP) and calcium oxide (CaO) and induces the crack and release of the coatings (Jeong et al., 2011; Surmenev, 2012; S Zhang et al., 2007). The tricalcium phosphate ($Ca_3(PO_4)_2$) is a CaP material with a dissolution rate greater than the hydroxyapatite and the CaO is a material that produce high toxicity levels when in contact with proteins and cell of the body.

On the other hand, to optimize the properties of osteoinduction, osteoconduction and osseointegration of the hydroxyapatite materials (in bulk) is necessary functionalize the chemical-structure with adequate quantities of ions substitutions as for example: Zn^{2+}, Sr^{2+}, Mn^{2+}, Mg^{2+}, Fe^{3+}, $(CO_3)^{2-}$, $(SiO_4)^{4-}$ and F^- by Ca^{2+}, $(PO_4)^{3-}$ and $(OH)^-$ (Miao et al., 2008; Pereiro et al., 2012; Surmeneva et al., 2013). In processes of coatings fabrication by chemical deposition is possible functionalize the HA structure with many other ions substitution, controlling mainly the stoichiometry and crystallinity but not the density, roughness, adhesion and thickness (Kim & Choe, 2011; Surmeneva et al., 2013). The physical deposition using plasma processes, produce a functional HA structure is not very common due to the complication of tune the plasma properties and control the deposition parameters to form this complex structure ($Ca_{10-m}M_m(PO_4)_{6-n}X_n(OH)_{2-q}Z_q$). In addition, several studies have shown that the Radio Frequency Magnetron Sputtering technique (RF-MS) have potential for applications in biomedical implant, that due to the use of a stable plasma along the time it could produce a nanostructured coatings with low surface roughness of high density and show good adhesion to the substrate (Nelea, Morosanu, Iliescu, & Mihailescu, 2003; Snyders et al., 2008; Snyders et al., 2007). Nevertheless, inadequate sputtering processes may produce low temperatures, energies and diffusion in the plasma that may inhibit the growth of a nanocoatings with a correct stoichiometry and high crystalline order (Kim & Choe, 2011; Nelea et al., 2003; Shi, Chen, Yu, & Zhang, 2008; Snyders et al., 2007; Van Der Wal et al., 2006).

In other works is possible observe that the thermodynamic properties of the plasma is the main responsible of coating formation (Han, 2009; Shaginyan, Han, Shaginyan, & Musil, 2006). About this, is very important understand know control and tuning the thermodynamic properties of the plasma for each plasma configuration or geometry. In this section is discussed the production of a high crystallinity and stoichiometric nanocoatings of hydroxyapatite and functionalized hydroxyapatite with fluorine ions obtained from the tuning of a highly magnetically confined plasma by orthopedic implant applications (Hong et al., 2007; López et al., 2013; López et al., 2015; Mello et al., 2007).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1.1. Deposition Techniques of CaP Coatings by Biomedical Applications

1.1.1. Plasma Spray

Plasma spray is the first commercial deposition technique to fabricate coating of CaP on metallic implants due the low cost and practicality. This technique consists in produce a plasma jet from a discharge at high frequency together with a gas injection (Khor et al., 2000). The gases commonly used in this process to produce the plasma jet are: the acetylene, butane and argon, which these gases limits the plasma temperature. After ignition of the plasma jet is aggregate the CaP powder. The plasma produces the ionization of the CaP particles that later are ejected to the substrate to form the coating. With this deposition technique is possible obtained CaP coatings with high hardness, low fracture strength, low coefficient of thermal expansion, high porosity, thick, low adhesion between coating and substrate (Khor et al., 2000). To produce HA coatings for this technique is commonly observed additional CaP phases as: TCP and CaO (Dong et al., 2003; Ellies et al., 1992). In addition, the crystallinity, density, thickness and roughness of the coatings are difficult to control producing not adequate mechanical properties[55] that in implant applications are not efficient.

1.1.2. Sol-Gel Deposition

This chemical deposition technique is extensively used for deposit CaP coatings over metallic implants mainly due to control of the HA stoichiometry and low cost (Guipont et al., 2010; Padilla et al., 2005; K. Zhang et al., 2005). The technique consist in polymerize an organometallic in an alcohol or water solution (Sol) together with an acid or base solvent (Gel) up to produce the bioceramic (Belabbes, de Carvalho, Schleife, & Bechstedt, 2011; Schwertfeger et al., 1994). With this technique is possible produce coating, fibers and powder (Schwertfeger et al., 1994). In coatings fabrication the process can form particles size of HA equal and less than the production by plasma spray (S Zhang et al., 2007). Commonly, deposition of HA and HA substituted with other ions may produce thick, porous and cracked coatings with weak adhesion and to form amorphous phases at room temperature (Chung et al., 2006; Jallot et al., 2005; Kim & Choe, 2011; Layrolle, Ito, & Tateishi, 1998; Tredwin, Young, Neel, Georgiou, & Knowles, 2014; S Zhang et al., 2007).

1.1.3. Electrodeposition

Several works has been performed using this technique to produce HA coatings. The mainly characteristics of these CaP coatings are low temperature processing, thick coatings ($\Lambda \approx 50$ μm), cracked, non-dense, high crystallinity, easy to coat complex parts, control of the composition and microstructure and long time to depositions. The technique involve the deposit over a metallic substrate by electrolysis controlling: the composition into the aqueous environment, the temperature, length between electrodes, voltage and current applied on the metallic substrate.(Dos Santos et al., 2010) From this technique is possible produce microcrystals of HA (5 μm of size) oriented along the *c* axis of the HA structure (Dos Santos et al., 2010; Zhang, 2013).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1.1.4. Pulsed Laser Deposition

The technique involve in ionize the particles of a target from the use of a pulsed laser of high energy density. The most used laser to deposited CaP films is the KrF ($\lambda = 248$ nm), (Carradò, 2011; Socol et al., 2004) (Iliescu, Nelea, Werckmann, & Mihailescu, 2004; Pelletier, Carradò, Faerber, & Mihailescu, 2011; Rau et al., 2010) but others are also used as: Nd:YaG ($\lambda = 355$ nm) and ArF ($\lambda = 193$ nm)(Jedynski, Hoffman, Mroz, & Szymanski, 2008; Mróz, Jedyński, Prokopiuk, Ślósarczyk, & Paszkiewicz, 2009; Saju, Reshmi, Jayadas, James, & Jayaraj, 2009).

When the laser beam of high energy strike the target surface the firsts micrometric layers may absorbed this energy producing a high atomic and molecular vibration, dissociating the atomic bond and consequently producing the ionization and ablations of atoms, radicals, clusters and agglomerates of particles that will be deposited over a substrate. The PLD technique may reproduce HA coatings with the same target stoichiometry, as well as, can reduce the thickness up to 50 nm. Commonly, these films produce a combination of crystalline and amorphous phases with high roughness and porous but with a better adhesion between substrate and coating (K. Zhang et al., 2005; Sam Zhang, 2013; S Zhang et al., 2007). In addition, is common perform a threat thermal *in-situ* on the substrate (TT ~ 350° C - 600° C --> 350 °C - 600 °C) to induce a formation of only crystalline HA phase (Surmenev, 2012; Surmenev et al., 2011).

1.1.5. Radio Frequency Magnetron Sputtering (RFMS)

This is the principal physical deposition technique to produce CaP nanocoatings (Akazawa & Ueno, 2013, 2014; Boyd, Duffy, McCann, Cairns, & Meenan, 2007; Feddes, Wolke, Jansen, & Vredenberg, 2003; Padilla et al., 2005; Snyders et al., 2007; Surmenev, 2012; Surmenev et al., 2011; Van Der Wal et al., 2006; Xu, Long, Sim, Diong, & Ostrikov, 2005; Yamaguchi, Tanaka, & Ide-Ektessabi, 2006). This technique consist in ionize an inert gas, commonly argon (Ar), producing a plasma which is confined by a magnetic field generated by the magnetrons. The Ar ions of the plasma are accelerated in direction to the cathode\target striking the target surface and consequently producing the pulverization of the target atoms and finally the deposition of these atoms over the substrate surface.

Most of coatings produced at room temperature by RFMS are composed by a lower crystalline order (some only of amorphous phase) and non-stoichiometric ($Ca/P \neq 1.67$). Threat thermal at $T > 500°$ C --> 500 °C (*in-situ* and *ex-situ*) are necessary to induce the crystallinity of the coatings and transform the ACP to HA coating (Boyd et al., 2007; Boyd et al., 2006; Ozeki, Aoki, & Masuzawa, 2010; Snyders et al., 2008; Snyders et al., 2007; Van Der Wal et al., 2006; Van der Wal, Wolke, Jansen, & Vredenberg, 2005). Nevertheless, in some cases the heat processes endanger the coating stability and the adhesion between coating and substrate.

On the other hand, all works using this technique has tried adjusting the deposition parameters for to get the fabrication of unique phase of crystalline and stoichiometric HA coatings. In this direction were produced and characterized CaP films using bias at high voltage and heaters at high temperatures on the substrate holder, which make the deposition system more expensive and difficult to operate. Despite the large number of publications on this theme, few studies has been carried out over the mechanisms of CaP films formation. Most of the research is limited to the knowledge of the relationship between the structural characteristics of the films with the deposition parameters such as: power density, voltage, gas pressure, and target - substrate distance, deposition time (rate), bias and temperature applied to the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

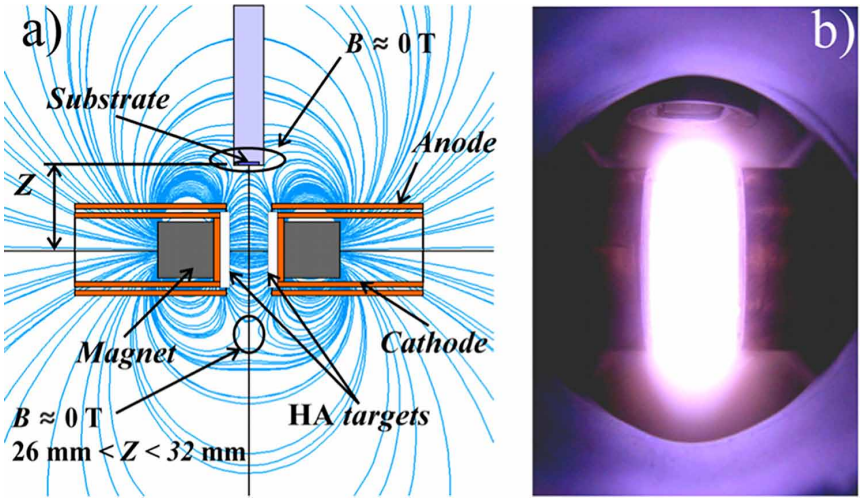
substrate (Surmenev, 2012; Surmenev et al., 2011; Surmeneva et al., 2013; Sam Zhang, 2013). In the case of physical vapor deposition techniques (PVD), have not still detailed studies involving the influence of the thermodynamic properties of the plasma with the films properties deposited on different substrates (Surmenev, 2012; Surmenev et al., 2011; Surmeneva et al., 2013).

1.2. Right Angle Magnetron Sputtering System (RAMS)

This sputtering system was adapted from the conventional magnetron sputtering geometry. This system exhibited two magnetrons positioned face-to-face (*in-axis*) with a substrate holder kept in a floating electric potential positioned in right angle to the magnetrons (*off-axis*). This modification was called right angle magnetron sputtering (RAMS) (López et al., 2013; Mello et al., 2007; López et al., 2015). The two magnetrons of the RAMS system work at 13.56 MHz of radio frequency (RF). The magnets into the magnetrons are composed for NbFeB (grade N42) and produce a high intensity magnetic field ($\mathbf{B} \approx 0.15$ T) which confined the plasma inside the central region of HA targets, as observed in Figure. 01 (López et al., 2013).

The high magnetic confinement of the plasma increases the ionization rate, the diffusion processes and the plasma temperature of the argon (Ar), oxygen (O_2) (5 mTorr and 1 mTorr, respectively) and the atoms that conform the targets (Ca, P and H). The substrates are positioned at distance of $\mathbf{Z} = 29$ mm from the center of the plasma bulk (see Figure. 01) which in this position there have a plasma trapped between magnetic mirror. The substrate holder was maintained electrically insulate with the RAMS chamber, producing a potential of 2 V relative to the ground, approximately. This floating potential and the deposition of ions (Ca^+, PO^-, O^-, H*) induces a great mobility of these species (ions and electrons) over the substrate surface, increasing the ions diffusion in the film growing as well as producing crystalline thin films conformed by HA nanodomains (in lower deposition time) that along the time the heat delivered by the plasma produces crystalline coatings of HA at room temperature (Figure 1).

*Figure 1. (a) Scheme of the RAMS system showing the magnetic field (B) and the substrate position region (**Z**) where the **B** ≈ 0 T. (b) Picture of the interior of the sputtering chamber showing the confined plasma glow at 120 W of RF power and at Ar and O_2 gas pressures of 5 mTorr and 1 mTorr, respectively López et al., 2013;*

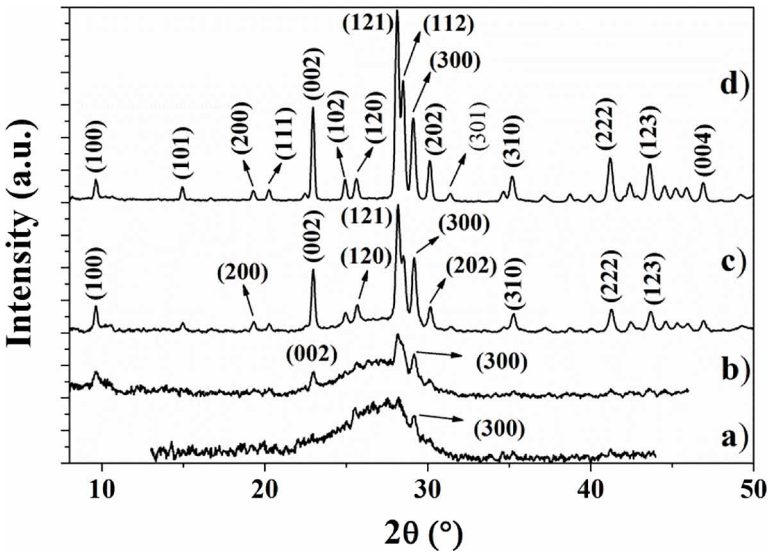


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1.3. HA Coatings Fabricated by RAMS System

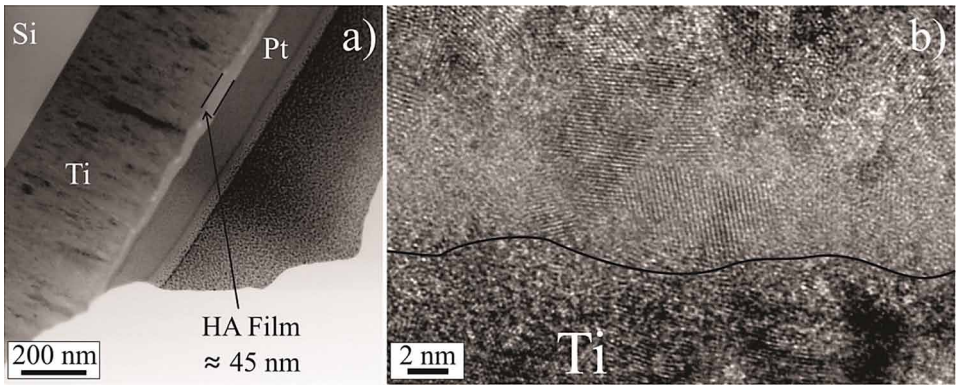
Commonly, there have a routine to coatings preparation which consist in associate the structural characterization (stoichiometry, crystallinity, adhesion film to substrate, roughness and texture) with the deposition conditions as: height between substrate - target, working pressure of argon and oxygen gases, bias potential on substrate holder, *in-situ* thermal treatment on substrates and RF power delivered to magnetrons (Surmenev, 2012; Sam Zhang, 2013). Other alternative is study, understand and control the thermodynamic properties of the plasma to form the coating with adequate physical-chemical properties and to propose new models of materials formations. For this, adequate knowledge in plasma technology and plasma processing must be taken in consideration as well as diagnose the plasma from adequate probes to measure the kinetic and dynamic of the electrons and ions in the discharge. The method most used to diagnose the plasma is from the Langmuir probe, which is possible characterize the thermodynamic properties of the electrons in the plasma discharge such as: density (n_e), temperature (T_{eff}), plasma potential (V_p), average kinetic energy ($< \varepsilon >$), diffusions (Γ) and energy delivered to film formation (E_D). (Chen & Smith, 1984; M. Lieberman & Lichtenberg, 2005; M. A. Lieberman & Lichtenberg, 1994; Shun'ko, 2009) With the RAMS system (using crystalline and stoichiometric and non-stoichiometric HA targets) was possible obtained stoichiometric (Ca/P = 1.68 ± 0.08) and crystalline (120 nm of crystallite size) nanostructured HA thin films (thickness less than 540 nm) with smooth roughness of 2 nm from the tuning of electron plasma properties in: $n_e = 2.0 \times 10^{15}$ m^{-3} and $T_{eff} = 11.8$ eV (Figure. 2 and Figure. 3) (López et al., 2013).

*Figure 2. GIXRD (λ = 0.13775 nm) patterns at a grazing incidence of $\Phi = 0.5°$ for the films that were fabricated with thicknesses and deposition times of (a) 90 nm – 30 min, (b) 135 nm – 45 min, (c) 225 nm – 75 min and (d) 540 nm – 180 min. The HA thin films were made at a deposition rate of 3 nm/min (**Z** = 29 mm) with a plasma power of 120 W and gas pressures of 5 mTorr and 1 mTorr for argon and oxygen*
López et al., 2013



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. (a) Image obtained by scanning transmission electron microscopy - STEM (bright field imaging) - for the sample that was prepared by FIB showing the CaP thin film region with a thickness of 45 nm between the Ti and Pt films. (b) HRTEM of the CaP thin film showing many HA crystalline nanodomains. The thin film was fabricated at 120 W of RF power, Z = 29 mm, deposition time of 15 min, deposition rate of 3 nm/min and gas pressures of 5 mTorr and 1 mTorr for argon and oxygen, respectively
López et al., 2013



From these results was possible to propose a growth model of stoichiometric and crystalline HA film produce at room temperature. This growth model associates the plasma temperature at $T_{eff} = 11.8$ eV and the total energy to form one HA unit cell, since ions deposition of Ca^{+}, O^{-}, H* and PO^{-} to formations of CaP aggregates, formations of CaP clusters precursors (possible Posner clusters), formation of a HA nanodomains and formation of crystalline films of HA (López et al., 2013). The energy of formation of these CaP clusters (precursors of the HA structure) were compared from a simulation made by *ab initio* in vacuum environment as proposed by other authors (Kanzaki, Treboux, Onuma, Tsutsumi, & Ito, 2001; Louis-Achille, De Windt, & Defranceschi, 2000; Treboux, Layrolle, Kanzaki, Onuma, & Ito, 2000).

The fabrication of a controlled crystallinity and stoichiometry of HA nanocoatings (at room temperature) produced without heat treatment (*in-situ* or after deposition) and bias is only possible from the knowledge and the control of the plasma properties, which will have a high efficiency in orthopedic implant applications.

1.4. Functionalized HA Coatings by Incorporation of Fluorine Ions and Fabricated by RAMS System

HA targets with ionic substitution of OH^{-} by F^{-}, ($Ca_{10}(PO_4)_6(OH)_{2-x}F_x$, where x = 0.6) was synthesized and produced to fabricated FHA films with the finality of functionalize the film surface and produce an increase of the physic-chemical stability, increase the resistance to dissolution by acid, reduce the bacterial proliferation and increase the bioactivity, as reported by other works (Louis-Achille et al., 2000; Montazeri et al., 2013; Sam Zhang, 2013).

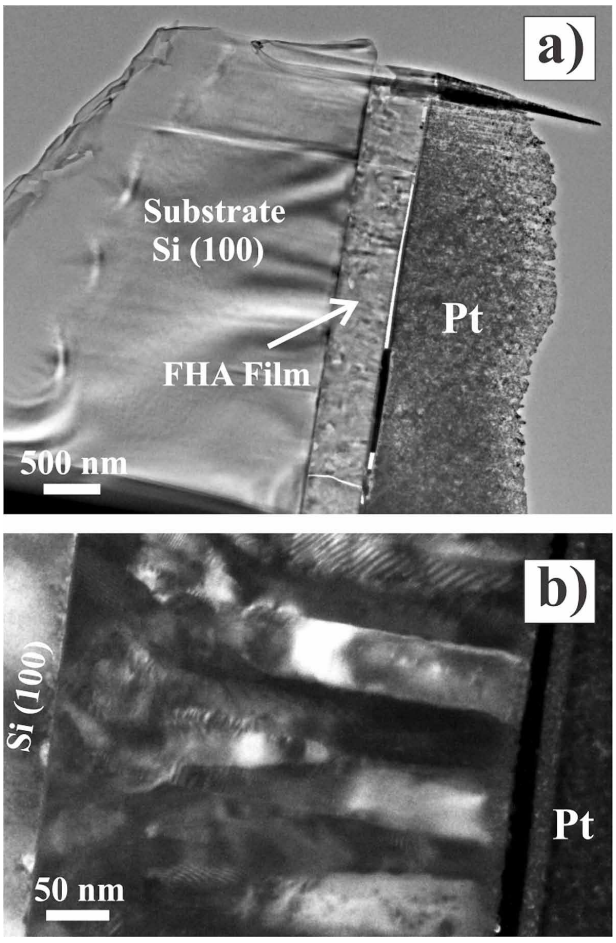
To produce FHA films with excellent crystallinity (in all film thickness as shown in the Figure. 04 and Figure. 05), excellent stoichiometry with adequate introduction of fluorine ions in the HA structure (Ca/P = 1.68 ± 0.04 and F/Ca = 0.12 ± 0.004), low roughness (R_{RMS} = 2 nm) and good reproducibility in different substrates surfaces, the deposition conditions was maintained at ***Z*** = 29 mm (height between

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

substrate and magnetrons, as observed in the Figure. 01), RF power of 110 W, deposition time of 180 min and deposition rate of 3 nm/min. In this condition the plasma parameters showed an electron density and temperature of: $n_e = 1.2 \times 10^{15}$ m^{-3} and $T_{eff} = 9.0$ eV, respectively (López et al., 2015).

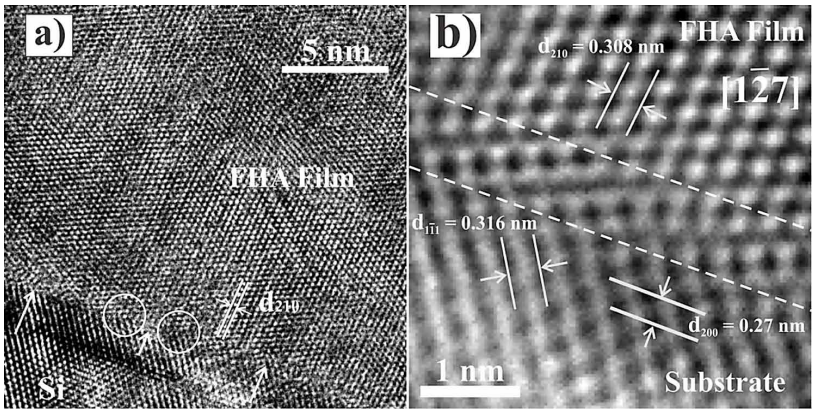
Finally, these results are evident that the right angle magnetron sputtering (RAMS) have a great potential to fabricate crystalline and stoichiometric nanocoatings ($\Lambda < 500$ nm of thickness) at room temperature of HA and FHA with adequate control of the fluorine incorporation and without the use of heater and bias on the substrate holder.

Figure 4. TEM images of 540 nm thick FHA film produced at 110 W RF power, 180 min and Z = 29 mm. The gas pressures of Ar and O_2 were 5.0 mTorr and 1.0 mTorr, respectively. (a) Low-magnification image of the sample prepared by FIB. (b) TEM dark field image showing columnar crystals oriented perpendicularly to the substrate surface
López et al., 2015



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. HRTEM (a) and filtered HRTEM (b) images showing the interface between an FHA film and a Si (100) substrate. The arrows in (a) and the segmented lines in (b) indicate the interface between the film and the substrate. The hkl planes are indicated in the images. The FHA film was produced at 110 W of RF power, 180 min and **Z** *= 29 mm. The gas pressures of Ar and O_2 were 5.0 mTorr and 1.0 mTorr, respectively*
López et al., 2015



2. ACOUSTIC WAVE NANO DEVICES ALUMINUM NITRIDE OBTAINED BY PULSED LASER DEPOSITION

Semiconductors nitrides of group III-V (GaN, InN AlN) alloys have been considered for many years as ideal for manufacturing semiconductor optoelectronic devices for the performance limit blue-ultraviolet (UV). The main appeal of these materials is the fact that these have a direct energy gap covering a wide range of energies from 1.9 eV (red) InN, to 3.4 eV (UVA) of GaN, to 6.2 eV (UVC) of AIN (Strite & Morkoç, 1992). Maruska *et al.* (Maruska & Tietjen, 1969) fabricate in the decade of the 60s the first GaN crystal thin film by epitaxy techniques hydride vapor phase. The main applications of these materials relate to emitters / detectors radiation devices in the blue-ultraviolet range. Today, most optoelectronic devices on the market operate at wavelengths from infrared (IR) to green.

2.1. Structural Properties of Nitride Semiconductors (Group III-V)

These nitrides crystallize according to three types of structure: Wurtzite ((hexagonal close packing)), also called (*α-AlN; α-GaN and α-InN*); zinc-blende (cubic), denoted as (β-*AlN,* β-*GaN and* β-*InN*) and rock salt structure ((Sodium chloride, NaCl Structure). The first two are the most common; the Wurtzite structure is the most stable at room temperature while the zinc-blende is metastable. The Wurtzite type structure consists of two hexagonal compact packages (HCP), two types of atoms interconnected along the axis direction of hexagonal symmetry, with a spacing of 5/8*c*. The Wurtzite zinc-blende structures are similar in that each metal atom (Group - III) is coordinated with four non-metallic (Nitrogen) and vice versa. In fact, they only differ in the sequence of stacking of planes diatomic N-Metal. Wurtzite structure in sequence in the direction <0001> is *ABABAB* while in the zinc-blende type *ABCABC* stacking is in the direction <111>. When having both structures similarity relates to appear zincblende type inclusions in layers with Wurtzite structure, especially in areas in which has stacking faults (Levinshtein, Rumyantsev,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

& Shur, 2001; Maruska & Tietjen, 1969; Morkoc, 2013; Ruterana, Albrecht, & Neugebauer, 2006). The lattice parameter of the three binary (GaN, AIN, InN) for the different structures is shown in Table 1.

The lattice vibrations (phonons), observable in the nitrides with Wurtzite structure by Raman spectroscopy, are related in Table 2.

2.2. Aluminum Nitride Applications in Acoustic Wave Sensors

The relationship between the surface acoustic wave (SAW) designed on thin films of aluminum nitride (AlN) grown on Si_3N_4 substrates by using a Pulsed Laser Deposition and optical properties has been widely studied due to their low net mismatch and thermal stability (Cressler, 2005). Moreover the dependency of both properties with a temperature of deposition. Aluminum nitride, an electrically insulating ceramic with a wide band gap of 6.3 eV, is a potentially useful dielectric material very important in fields such as optoelectronic and microelectronics. For instance, it can serve as the dielectric gate with high voltage and high power electronic devices; AlN is also being investigated as a substitute for the silicon dioxide buried layer in silicon-on-insulator (SOI) substrates due to its low thermal expansion coefficient, high breakdown dielectric strength, and high chemical and thermal stability (Clement, Vergara, Sangrador, Iborra, & Sanz-Hervás, 2004; Jiang, Shen, Ogawa, & Guo, 2003; Yamashita, Fukui, Misawa, & Yoshida, 1979). Moreover, metal nitrides based on AlN are of high interest because of their electrical, optical and acoustical properties (Clement et al., 2004). Their hardness and thermal coefficients of expansion

Table 1. Lattice parameters and energy gap of nitrides of group III-V at room temperature

	Lattice Parameters (Å)	**EG (300K) (eV)**
β-AlN	$a = 4.38$	5.11 (theoretical)
α-AlN	$a = 3.112$ $c = 4.982$	6.2
β-GaN	*a = 4.52*	3.2
α-GaN	$a = 3.189$ $c = 5.185$	3.39
β-InN	*a = 4.98*	2.2 (theoretical)
α-InN	$a = 3.548$ $c = 5.705$	1.89, 0.9

Levinshtein et al., 2001; Morkoc, 2013; Ruterana et al., 2006

Table 2. Nitrides vibrational modes in Wurtzite structure. LO: Longitudinal Optical; TO: Optical Transversal

Vibration Mode (cm^{-1})	**AlN**	**GaN**	**InN**
E2 – Low frequency	215	137	85
E2 – High frequency	652	592	485
A1 – *TO*	640	546	449
E1 – *LO*	883	732	587
E1 – *TO*	668	555	457
E1 – *LO*	922	741	596

Levinshtein et al., 2001; Morkoc, 2013; Ruterana et al., 2006

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

are comparable to that of Si. AlN is used for acoustic wave devices on Si (Clement et al., 2004), optical coatings for spacecraft components, heat sinks in electronic packaging applications, as well as electroluminescent devices in the wavelength range from 215 nm to the blue end of the optical spectrum. Some optical properties near the fundamental band gap of AlN have been reported by Yamashita et al. (Yamashita et al., 1979) since 1979.

Several theoretical studies have been well documented in the past, (Jiang et al., 2003; Levinshtein et al., 2001) however, reliable experimental data has been reported just recently because both the sample quality and the spectroscopic techniques have only been substantially improved in recent years. Precise knowledge of the optical constants is particularly important in view of the use of AlN thin films in optical filters and light emitting laser diodes since they also show temperature dependence. Similar to the section 1.1.4 for the HA, Pulsed Laser Deposition (PLD) is largely applied for processing thin films and other structures. The highly no equilibrium nature of the PLD process is attractive for the synthesis of stoichiometries thin films of various metal nitrides, carbon nitrides, oxides and oxynitrides from the corresponding bulk targets. PLD appears to be a suitable method to transfer stoichiometrically complex monolayer structures from AlN ceramic structures to substrates.

Thus, from AlN stoichiometries films, many infrared (IR) measurements presented in the literature were used to estimate the film quality. The IR-spectra provided important information about the composition, homogeneity, crystallinity and the residual stresses present in the films. Moreover, the ellipsometry and reflectance measurements have been used to determine the refractive index in many AlN compounds, since the ellipsometry and reflectance are sensitive and non-destructive techniques used for studies of optical properties and microstructures of surfaces and thin films (Caicedo, Pérez, Caicedo, & Riascos, 2013; Pérez, Riascos, Caicedo, Cabrera, & Yate, 2011a; Rosenberger, Baird, McCullen, Auner, & Shreve, 2008; Zhuang et al., 2009). Additionally, single-chip front-end RF modules incorporating Surface Acoustic Wave (SAW) filters are a matter of intense research. At present, these modules are fabricated by bonding single crystal piezoelectric SAW devices onto integrated circuits. Thin films of polycrystalline Aluminum Nitride (AlN) are promising materials for the integration of SAW devices on Si_3N_4 substrates due to their good piezoelectric properties and the possibility of deposition a low temperature compatible with the manufacturing of Si integrated circuits. AlN thin films of sufficient quality for SAW applications can be a obtained by the PLD technique (Pérez et al., 2011a).

2.3 X-Ray Photoelectron Spectroscopy (XPS) Analysis

The XPS survey spectrum of AlN thin film grown at 300 ºC is shown in Figure 6. The peaks at 534.4 eV, 397.3 eV, 124.4 eV, and 73.9 eV correspond to O1s, N1s, Al2s, and Al2p binding energies, respectively. Calculation of the peak areas of deconvoluted spectrum (Figure 7) without O1s contribution gives an atomic ratio of Al:N = 0.392: 0.588, which is similar to the stoichiometry of $Al_{0.40}N_{0.60}$ (Caicedo et al., 2013). The core electronic spectra carry the information about the chemical composition and bonding character of the Al–N films. The integral of N1s, Al2s, and Al2p spectra corrected by relevant sensitive factors can evaluate the concentrations of Al, and N elements in the Al–N films. The corresponding integral of the deconvoluted peaks can also be used to estimate the bond contents, which are described by the following formula:

$$C_i = \Sigma(A_i/S_i)/\Sigma(A_j/S_j) \quad (1)$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

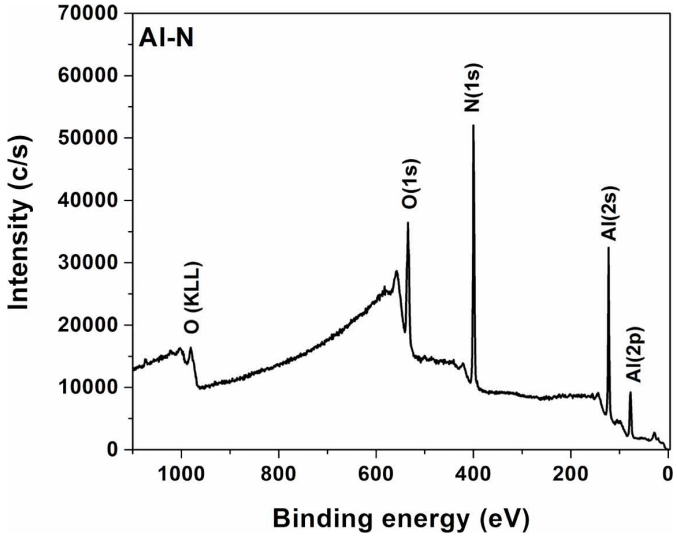
Where **-*S*** is the sensitivity factor, ***A*** is the integral of deconvoluted peaks, and ***C*** is the atomic content. The numerator is the sum of the integral of one sort of bond; the denominator is the sum of the integral of all types of bonds decomposed from the whole peak of N1s, A2s, and Al2p in the sample. The atomic concentrations of Al, N and O elements of XPS analysis of the coating deposited at 300 °C are listed in Table 3. For all the Al-N films obtained, the O concentration is less than 3 at.%. We can speculate that the majority of the oxygen at the surface of the non-treated sample (which is smooth and almost perfect) is chemically bonded into alumina-dominated oxide- nitride material.

The high-resolution spectra of Al2p, and N1s were recorded from AlN films, as shown in Figure 7(a), the Al2p peak is composed of a shoulder separated by 1.7 eV with intense peak. The XPS spectrum of Al2p can be fitted well by two Gaussian functions. The value of the binding energy obtained for the Al2p peak was 73.9 eV and the higher value for Al2p was 75.6 eV, respectively. According to the literature(Mahmood et al., 2003; Pérez Taborda, Riascos Landázuri, Jiménez García, & Caicedo Angulo, 2010; Taborda, Caicedo, Grisales, Saldarriaga, & Riascos, 2015) for the Al2p peak, the first one (73.9 eV) and the second one (75.6 eV) can be assigned to Al–N and Al-O bonds.

The appearance of the peak at 73.9 eV clearly shows that Al has reacted with N; therefore, it can be assigned to Al-N. In Figure 7(b), N1s peak is composed of spin doublets, each separated by 2.9 eV. The XPS spectrum of N1s, well fitted by two Gaussian functions, depicts the N1s spectrum with values at 397.3 eV and 400.2 eV characteristics for N-N and Al-N bonds, respectively.

The XPS results demonstrated that Al atoms bonded to N in the form of nitride, taking into account that the elemental concentration of the Al-N film was obtained by adjusting the laser incidence on Al target, and the N_2 as a working gas. In this specific case, it was found that amounts of Al (N) in the AlN film were maximum at the current establishment of process conditions and the ratio of Al and N in the film was about for 2:1. In general, formative Al (N) phase indicates that the aluminum and nitrogen activity and activation energy provided by the present deposition conditions are enough for the formation of an Al-N thin film. Although the surface temperature of the substrate during deposition of Al-N

Figure 6. XPS survey spectrum of Al-N coatings deposited on Si_3N_4 at 300 °C, where different elemental signals show



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

film is around 300 °C, the substrate lies in a high-density plasma region and a high ion-to-atom ratio of aluminum and nitrogen, thus behavior can induce to the formation of Al-N phase at the low temperature below 300 °C. Moreover, the correlation between temperature deposition and chemical properties due to the reduction of oxygen concentration when the increase of deposition temperature, this behavior was associated with high surface energy activity generated by high surface temperatures that induce more nitrogen reaction.

The concentrations in these depth profiles were obtained by curve fitting of high-resolution XPS spectra and measurements of the XPS peak areas. In order to see the effects of increasing the substrate temperature during the growth of thin films which would serve AlN buffer for Mo SAWs. The dependence of aluminum and nitrogen in the deposition temperature is shown in Figure 8 which represent various chemical states as described in the section 2.3 XPS in Table 3 and Figure 8.

Figure 7. High-resolution XPS spectrum of: (a) Al2p and (b) N1s of Al-N film is around 300 °C

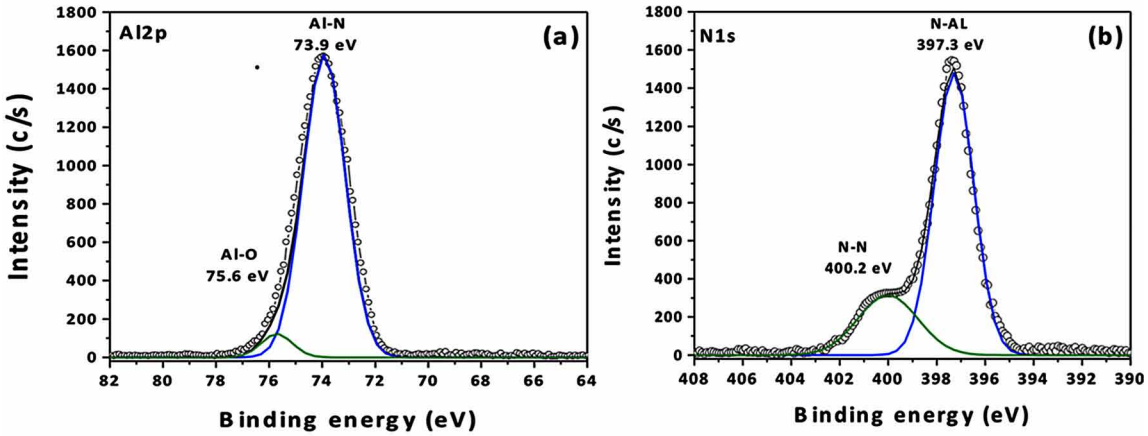


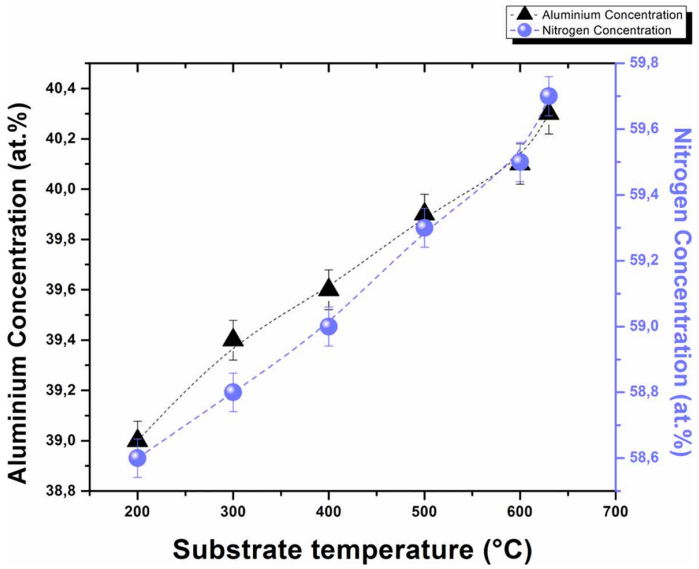
Table 3. XPS analysis-derived atomic concentrations of Al, and N elements in the Al-N film synthesized by PLD with 7 mbar and substrate deposition temperatures as function of atomic percentage (at%)

Growing Conditions		Atomic Composition (at %) Al-N Films		
Substrate Temperature (°C)	Pressure (10^{-3} mbar)	Al	N	O
200	9	39.0	58.6	2.4
300	9	39.4	58.8	2.0
400	9	39.6	59.0	1.8
500	9	39.9	59.3	1.5
600	9	40.1	59.5	1.3
630	9	40.3	59.7	1.1

(Caicedo et al., 2013; Pérez, Riascos, Caicedo, Cabrera, & Yate, 2011b; PEREZ TABORDA, RIASCOS, & VERA; Taborda et al., 2015)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Temperature dependence at the concentrations of Al and N for AlN films



2.4 Surface Analysis of AlN Films via Atomic Force Microscope

Nevertheless, most of the energy in the propagating surface acoustic waves is contained in the region within one wavelength below the surface, so they are very sensitive to the surface state and are affected by any change, such as residual stresses and the level of roughness. Moreover, authors such as C. Zhuang *et al.(Zhuang et al., 2009)* have shown that at the same frequency and in the small roughness limits (RMS ≪ wavelength), the SAW velocity increases according to the decrease in the RMS roughness. Figure 9 a-c shows a 5μm × 5μm image for AlN films deposited with different deposition temperatures. The general trend in the AlN films with the decreasing temperature exhibits a decrease of the grain size and roughness. In Figure 9 obtained from AFM drops observed of which are characteristic of the PLD technique and are occasioned by the release of particulates from the Aluminum target at impact with the substrate, such generate imperfections on the surface of the films also shows micro drops on AlN surface generated by PLD energy deposition corresponding to high laser fluency. Moreover, the correlation between roughness and grain size with deposition temperature is shown in Figure 10 The quantitative values were extracted from the AFM images by means of statistical analysis Scanning Probe Image Processor (SPIP®). Each data point in the plots represents an average over AFM images. The corresponding error bar was obtained by a standard deviation of the values (Mahmood et al., 2003; PEREZ TABORDA et al.; Pérez Taborda et al., 2010; Taborda et al., 2015; Zhuang et al., 2009).

Figure 10 shows the grain size and roughness values for the AlN films, thus, in this case, a decreasing tendency in surface measurements was observed when the deposition temperature was increased. The lower grain size and roughness values obtained in this study by the Al–N films deposited on silicon substrate represented a decrease of grain size around 42% and roughly 31%, compared to that obtained from the Al–N films deposited at 200 ºC. This fact is relevant since the continuity of the surface is important for optical and electrical SAW devices (Kar & Bose, 2011; Raole, Mukherjee, & John, 2005).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. AFM images for AlN films grown onto Si_3N_4 substrate with different deposition temperatures: (a) 200 ºC, (b) 400 ºC and (c) 630 ºC.(Caicedo et al., 2013)

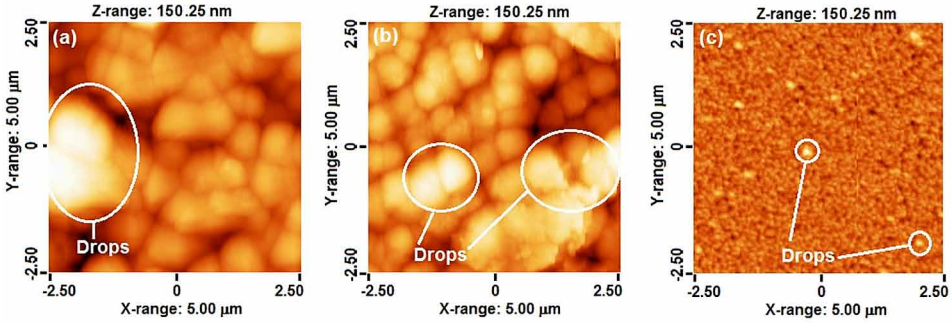
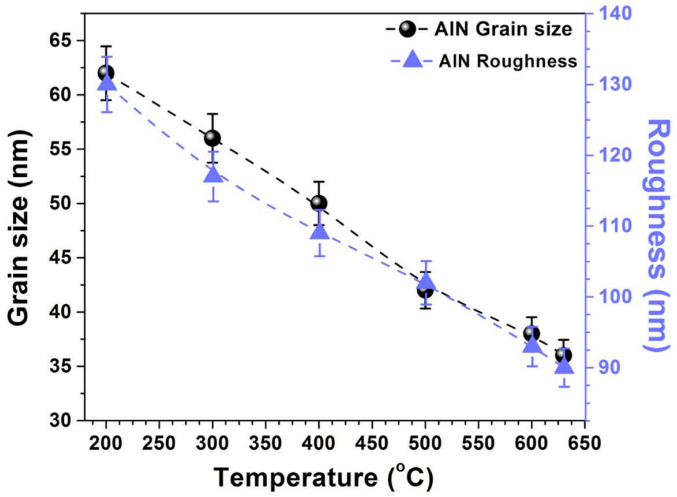


Figure 10. Surface measurements obtained via AFM and SPIP® analysis for AlN films with a total thickness of 150 nm: Correlation between grain size and roughness with deposition temperature



The variation of grain size and roughness (Figure 10) obtained in the Al–N films suggest that thermal activation generated by the increasing in the deposition temperature on surface substrates modifies the superficial morphology by increasing the energy associated with atoms on the substrate surface and/or the growing film surface which produces a relative reduction of grain size and roughness.

2.5 Dominant Wavelength and Color Purity Analysis

The reflectivity measure is the fractional amplitude of the reflected electromagnetic field while reflectance refers to the fraction of the incident electromagnetic power that is reflected at an interface. The reflectance is thus the square of the magnitude of the reflectivity. The reflectivity can be expressed as a complex number as determined by the Fresnel equations for a single layer, whereas the reflectance is always a positive real number.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Additionally, reflectivity is distinguished from reflectance by the fact that reflectivity is a value that applies to thick reflecting objects. When reflection occurs from thin layers of material, internal reflection effects can cause the reflectance to vary with surface thickness. Reflectivity is the limit value of reflectance as the surface becomes thick; it is the intrinsic reflectance of the surface, hence irrespective of other parameters such as the reflectance of the rear surface (Kar & Bose, 2011; Touati, Hadjoub, & Doghmane, 2006; Zayats et al., 2008). On the other hand, the dominant wavelength (of a color stimulus) is defined as: "*the wavelength of the monochromatic stimulus that, when additively mixed in suitable proportions with the specified achromatic stimulus, matches the color stimulus considered*" (Kar & Bose, 2011). The AlN films have been compared with pure aluminum because the mirror finish aluminum has the highest reflectance of any metal in the 200–500 nm and the 3000 –10000 nm (far IR) regions, while in the 500 –700 nm visible range it is slightly outdone by AlN and silver, but in the 700 –3000 nm range (near IR) it is slightly outdone by, gold, and copper materials.

The reflectance spectrum or spectral reflectance curve is the plot of the reflectivity as a function of wavelength as can be seen in Figure 11 Therefore, all the deposited films appear highly reflective with white and aluminum colored. Figure 11a. shows the optical reflectance spectra of the AlN single layers obtained at different deposition temperatures. The reflectance of the aluminum and the eye sensibility are shown for comparison. The spectra of the samples show high reflectance for long wavelengths, near to 33% for the AlN films deposited with 200 °C and close to 23% of AlN films deposited with 630 °C. These values of reflectance at these wavenumbers are in good agreement with previous reports for AlN films (Zhuang et al., 2009). A clear decrease in reflectivity for short wavelengths is seen, characteristic of a system with high free-electron density with a reflectance edge below 550 nm due to a screened plasma resonance. The white and aluminum colors of the AlN films are a result of the steep plasma reflection edge that occurs in the visible region where the reflectivity minimum is around 560 nm. The color purity refers to the degree of purity present at a color; therefore in a pure color do not exist any achromatic effect within its composition, which can vary its capacity for reflection. One of the three dimensions of color refers to the color purity that a surface can reflect. When the tone is full and is not mixed with an achromatic effect (brightness), this presents its maximum saturation or purity. On the other hand, if a particular color is combined with an achromatic effect, or with its complement (luminosity), it suffers a loss of purity or neutralization. The color purity changes with the temperature of deposition, at ranges from 0.63 to 0.66 for AlN films, away from the color purity of pure aluminum (0.90), hence confirming that the films are less white compared with the aluminum. Color coordinates of AlN single layer films are shown in Figure 11b.

Reference coordinates of pure aluminum are shown for comparison in that Figure 12. The results of color measurement indicate that all films reflect a hue slightly shifted far to the white as compared to aluminum reflectivity and with similar color purity. Likewise, dependence is observed the high dependence on reflectance percentages when the deposition temperature is varied from 200 to 630 °C. Figure 12a exhibits one constant region for wavelengths ranged from 760 to 800 nm; in those regions it is possible to appreciate the effect of deposition temperature on changes in the reflectance values. Figure 12b shows the decreasing of reflectance when the temperature is increased, which indicate that temperature promotes the absorbance in the AlN deposited via PLD. The last behavior changes in optical properties (reflectance) can be related with changes generated by the reduction of grain size and reduction of surface roughness when the depositing temperature is increased producing an activation surface energy observed from AFM images.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Dominant wavelength and color purity results (color online): (a) Optical reflectance of AlN films deposited onto Si_3N_4 (100) substrates at the different deposition temperature, also aluminum optical reflectance and eye sensibility were also plotted as the references. (b) Chromatic diagram, in the x, y coordinates, of the reflectivity for AlN films. White coordinates of an achromatic point are located at (1/3, 1/3)

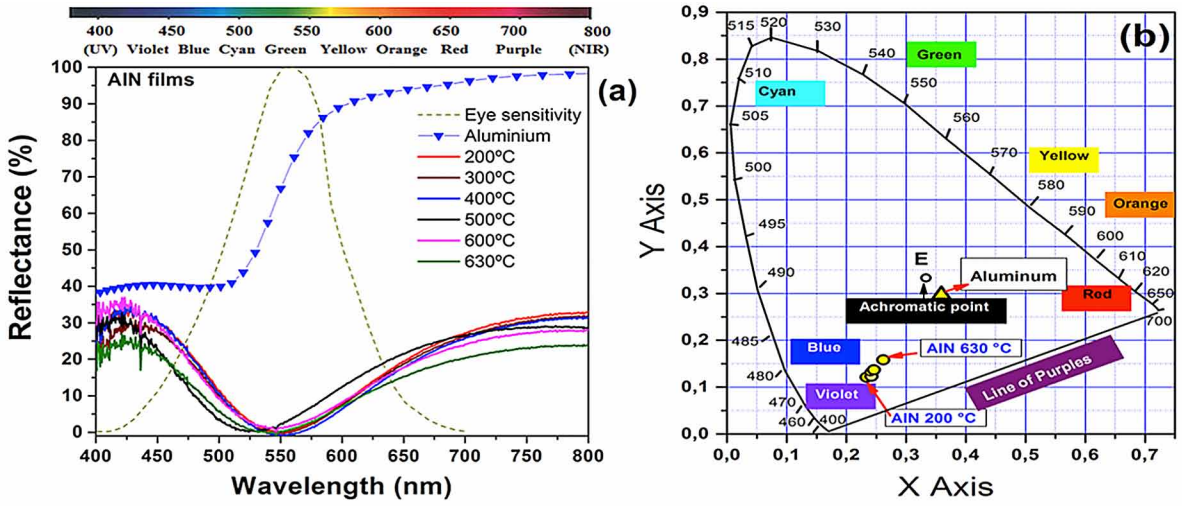
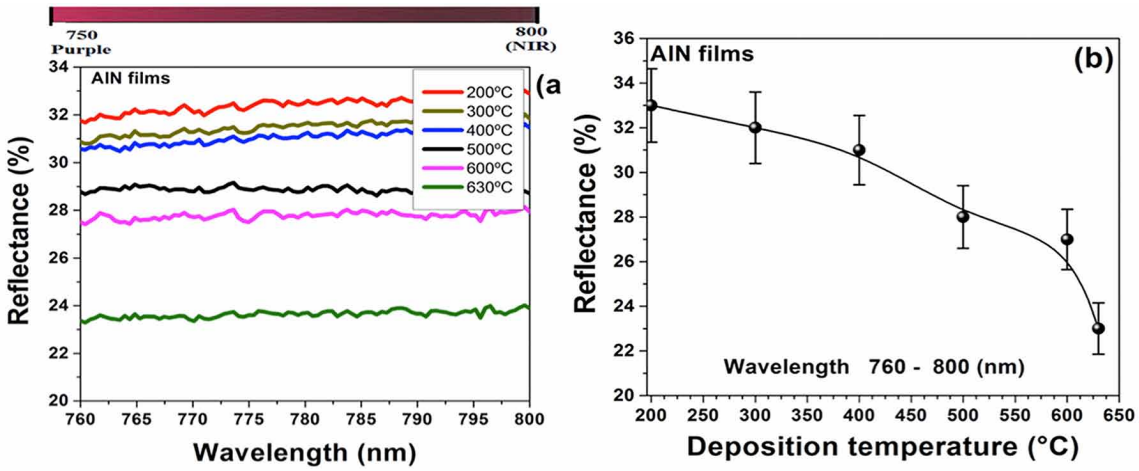


Figure 12. Temperature-Reflectance dependence (color online): (a) Reflectances of AlN films deposited onto Si_3N_4 (100) substrates at the different deposition temperature from 760 to 800 nm wavelength, (b) reflectance as function of deposition temperature



From the reflectance spectra it is seen a weak but clear effect of deposition temperature on the optical properties. As the temperature is decreased, the reflectance of the films tends to be higher in the near infrared region. So, dominant wavelength and color purity have been calculated from the reflectance spectra of all films by using the *java-optics* software (Table 4). The dominant wavelength was varied for all the samples, from 562 to 570 nm, and a little close to that of pure aluminum reference (574 nm)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In. Figure 13a it can be observed the influence of deposition temperature on color purity. This graph shows the increase of purity values towards color gray purity. Figure 13b shows the differences in the dominant wavelength for all AlN films deposited with different deposition temperatures.

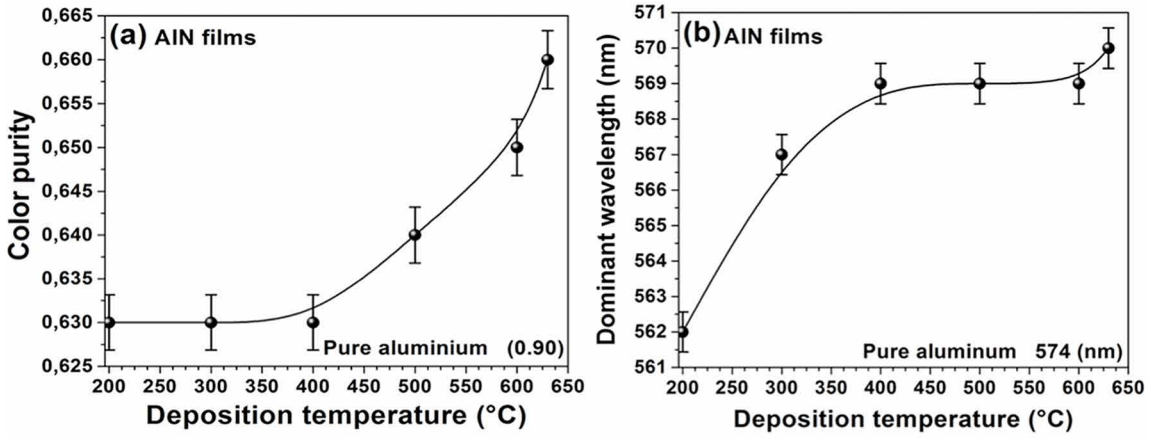
The wavelength is an important optical characteristic for different materials. When the wavelength is changed it is possible to observe that natural color is changed. This modification on optical properties is associated to changes generated by the reduction of grain size and reduction of surface roughness observed in AFM images. Moreover, the changes in chemical composition of AlN films observed in XPS spectra show that oxygen concentration is reduced when the temperature deposition is increased. Therefore, the low oxygen concentration induce changes in the energy bands generating changes at optical properties (color purity and dominant wavelength). Finally, the purity color dependence in AlN films with temperature obtained demonstrates the possibility of some purity color control.

SEM images was carried out to quantitatively study the surface morphology of our samples in relation to an increase on deposition temperature in AlN films grown onto Si_3N_4 (100). Figure 14 Shows SEM micrographs for AlN films with a random distribution of microparticles or micro drops that were analyzed on these surfaces in agreement with AFM images. Accordingly, the deposition temperature

Table 4. Optical properties for AlN films deposited onto a Si_3N_4 substrate

Deposition Temperature (°C)	Dominant Wavelength (nm)	Color Purity	Axis X	Axis Y
200	562	0.63	0.234	0.121
300	567	0.63	0.242	0.123
400	569	0.63	0.240	0.132
500	569	0.64	0.246	0.137
600	569	0.65	0.262	0.158
630	570	0.66	0.302	0.270
Reference Aluminum	574	0.90	0.358	0.371

Figure 13. Temperature dependence at color results for AlN films: (a) color purity, and (b) dominant wavelength as a function of deposition temperature



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

clearly affects the reduction of micro-drops; this can be possible due to great surface mobility when the temperature was varied from 200 to 630 °C. This surface mobility reduces the possibility that the micro-drops are anchored on the surface when arriving with high energy on AlN film. Otherwise, the high superficial temperature that produces surface diffusion of micro-particles or micro-drops, which can decrease the overall number of particles. The magnification image showed in Figure 15 exhibits that molybdenum lines are continuous, without any interruption or short circuit. After electrical measurements, it was possible to deduce an insulator behavior for our AlN structures, a good prerequisite for using them as surface acoustic wave (Mo/AlN/SAW) devices.

2.6 Applications: Frequency response of Mo/AlN SAW devices

Authors as Hadjoub *et al.* (Hadjoub, Touati, Doghmane, & Doghmane, 2008) been reported improvement in the Rayleigh Velocity (VR) by decreasing the thickness of the films and substrates used Si_3N_4 with a velocity of 5694 m/s value very next to that of VR (SiC) = 6810 m/s. These values are referred to substrates for fast VR (Al_2O_3) = 5680 m/s, VR (Si) = 4710 m/s, VR (SiO_2) = 3410 m / s, VR (Ti) =

Figure 14. SEM micrographs where is observed the micro-particles or micro-drops: (a) AlN films deposited with 200 °C, (b) AlN films deposited with 630 °C

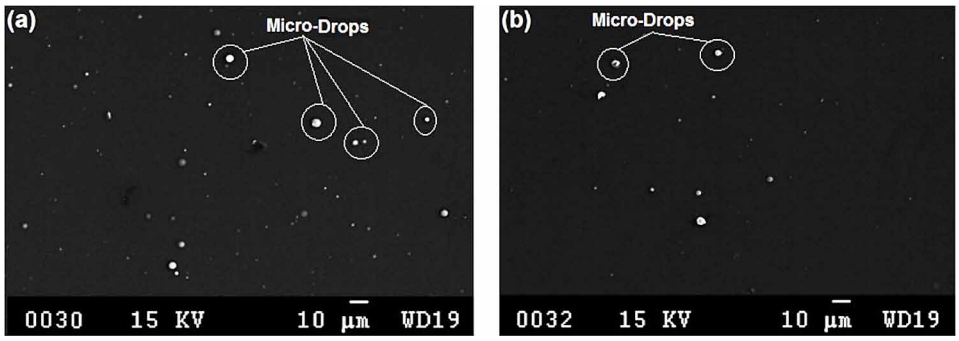
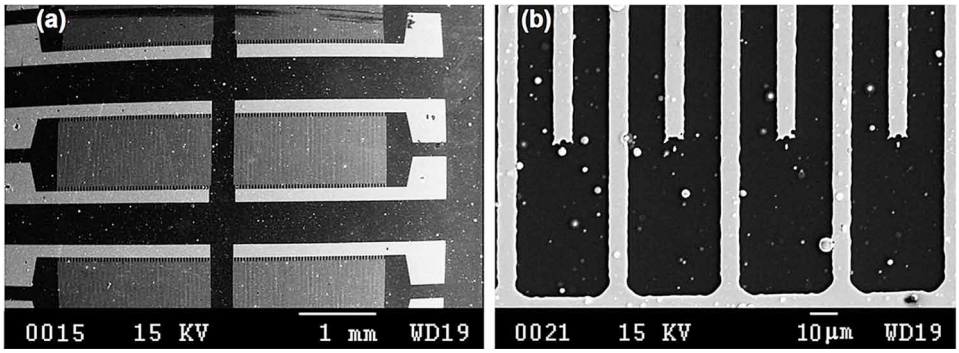


Figure 15. SEM micrographs showing the interdigital structure of Mo coating on AlN films obtained by PLD: (a) general interdigital structure, (b) magnification images where it is possible to observe the width and distribution of lines that conform the interdigital device



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2960 m/s and VR (W) = 2670 m/s, reported in the literature (Caicedo, Zambrano, Aperador, Escobar-Alarcon, & Camps, 2011; Drafts, 2001; Hashimoto & Yamaguchi, 2001; Kao, Cheng, & Chen, 2002; Paduano, Weyburne, Jasinski, & Liliental-Weber, 2004). It clearly shows the importance of using substrates with good speed of Rayleigh and compatible with existing silicon technologies so that in our case Si_3N_4 because VR is 20% higher than that of Si.

Recently, some authors (Caicedo, Zambrano, Aperador, Escobar- Alarcon, & Camps, 2011; Drafts, 2001; Hashimoto & Yamaguchi, 2001; Kao, Cheng, & Chen, 2002; Paduano, Weyburne, Jasinski, & Liliental-Weber, 2004) have reported the relation between the h/λ and V_{SAW} by experiment or theory, where h is AlN thickness, λ is wavelength, and V_{SAW} is SAW velocity. The relationship between the thickness and the wavelength this $h/\lambda = 0.004$. The frequency response characteristics of Mo/AlN SAW devices fabricated using the different deposition conditions, such as, 200 ºC, 400 ºC and 630 ºC, have been measured, of which results are shown in Figure 16 a-d. As can be clearly seen by comparing the wide-frequency scan spectra of those three devices (which are plotted in the insets of Figure 16c, by using the AlN buffer layers the center frequency was shifted from approximately 0.618 GHz Figure 16a to higher frequencies of 0.621 GHz (Figure 16b) and 0.625 GHz (Figure 16c). This indicates that the AlN buffer layer could be used as high phase velocity substrate in the fabrication of high-frequency SAW devices, as with Molybdenum.

Additionally, in Figure 16d that the insertion loss was reduced by the use different deposition temperatures of AlN buffer, by using the AlN buffer layers the center frequency was shifted from approximately 0.618 GHz (Figure 16a) to higher frequencies of 0.621GHz (Figure 16b) and 0.625GHz (Figure 16c). This indicates that the AlN buffer layer could be used as high phase velocity substrate in the fabrication of high-frequency SAW devices. In addition, it appeared in Figure 16c that the insertion loss was reduced by the use of AlN buffer. To examine in more detail the dependence of insertion loss on the deposition condition of AlN-buffer, the insertion loss has been measured for all the fabricated SAW devices.

The results are plotted in Figure 15 as a function of temperature deposition of AlN films and the estimated quantitative values are also listed in Figure 16. As can be seen in Figure 17, for SAW devices, fabricated using the sputtered-AlN buffer with Mo channel the insertion loss was exactly correlated with the corresponding temperature value for all AlN films.

This can be identified by comparing the results obtained from the device with AlN films deposited on silicon nitride substrate. For the above devices, the temperature deposition was increasing systematically, as already shown in Figure 17. Moreover, the device with AlN deposited to 630 ºC revealed the relatively large insertion loss in relation to frequency value (0.625 GHz). This may be attributed to the relatively low surface roughness (90 nm) AlN film used in Mo/AlN device considering that the surface roughness of AlN film may also be associated by the surface morphology and deposition temperature of AlN buffer layer, the results shown in Figure 16d. confirm that frequency response characteristics (insertion loss) can be affect by reduction of oxygen concentration, reduction of grain size, roughness, and decreasing with reflectance of AlN films together with increase of color purity and dominant wavelength thin this sense, the property of AlN buffer may strongly govern the characteristic of SAW device using the Mo/AlN system. (Caicedo et al., 2013)

The direct relationship between increased frequency and speed in SAW devices, all this related to the morphological qualities of roughness and grain size favor in increased rate of velocity. The resonance frequency of an SAW is determined by the equation $f = V/k$, where f is the resonance frequency, V is the phase velocity of the acoustic wave and k is the acoustic wavelength, which in turn is defined by the electrode pitch in the *interdigital transducer* (IDT). The reflected signal of the Rayleigh mode in AlN

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. Frequency response characteristics of fabricated Mo/AlN SAW devices with different deposition temperatures (a) 200 ºC, (b) 400 ºC, (c) 630 ºC and the insertion loss by the use different deposition temperatures of AlN buffer for (a), (b) and (c)

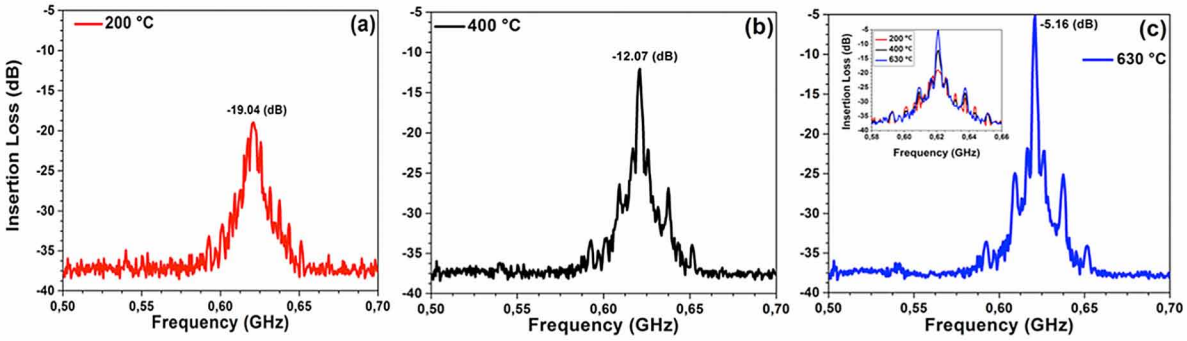
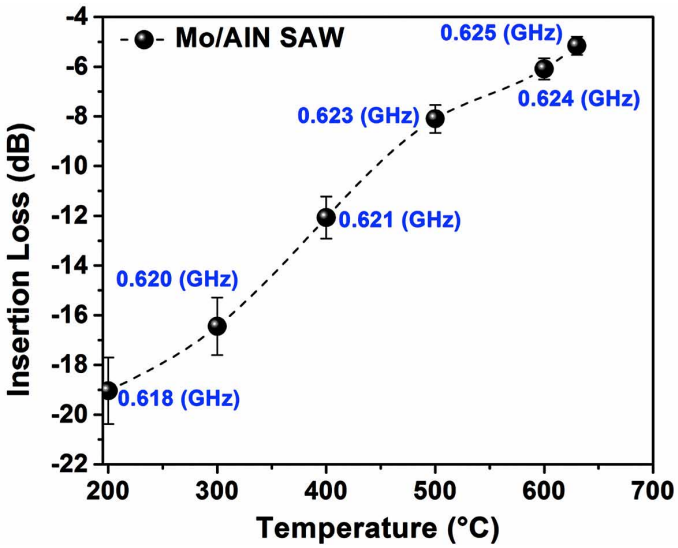


Figure 17. The measured insertion losses of fabricated Mo/AlN SAW devices, as a function of the temperature values from the deposited AlN films

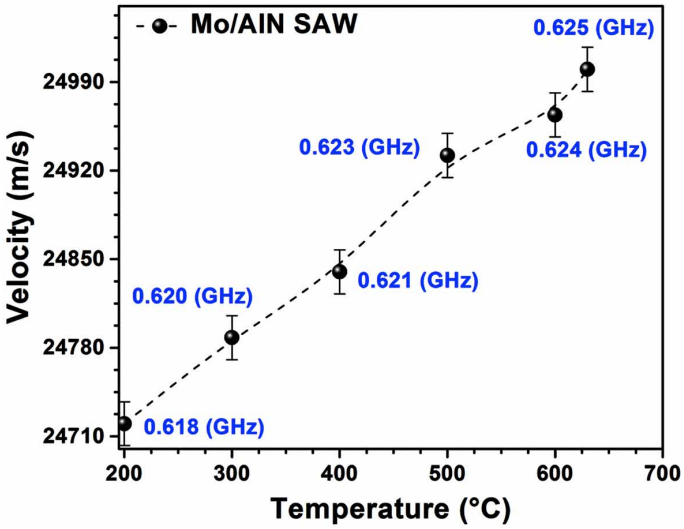


films deposited to highest temperatures (630ºC) is observed to be much stronger than that of the AlN films deposited to lowest temperatures (200 ºC) and the phase velocity of AlN films deposited to (630 ºC) is much higher. For The Mo/AlN SAW devices, the measured resonant frequencies of the Rayleigh mode waves ranges from 0.618 GHz to 0.625 GHz corresponding to phase velocities of 24720 m/s to 25000 m/s, indicating good repeatability and reproducibility (Pérez et al., 2011a).

Thus, the lowers deposition temperature of the substrate, generate a perturbation of the acoustic velocity that is related to a change in frequency of the oscillator circuit. In the Figure 17 is shown as before said to improve velocity with increased temperature of the films deposited during the growth of AlN and were used as substrates for SAW devices Mo. Figure 18 Acoustic wave velocity as a function of deposition temperature in Mo/AlN/Si_3N_4 system.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. Acoustic wave velocity as a function of deposition temperature in Mo/AlN/ Si3N4 system



Take the last behavior is possible to observe that the quasi-static electric field associated with the SAW interacts with the charge carriers of the film. The SAW velocity is, therefore, dependent on the AlN thin film conductivity changes. These velocity changes can easily be interpreted as a shift in the resonant frequency, when the SAW delay line is incorporated in an oscillator conFigureuration. The effect of temperature is one of the easiest parameters that can be sensed because almost all physical and chemical effects depend on temperature (Kar & Bose, 2011; Morkoc, 2013).

In certain applications, the effect of temperature should be minimized in order to avoid any interference with other effects. For acoustic wave sensors, the temperature effect induces dimensional dilatations and small changes in the mechanical and electrical properties. Two kinds of interaction can thus be measured: *(i)* shift of the acoustic wave velocity due to intrinsic changes of material properties (elasticity and density) (this is the case for almost all SAW devices using metal-nitrides materials with low thickness) (Aissa et al., 2015; Caicedo et al., 2013; Clement et al., 2004; Jung, Lee, Kim, & Park, 2004; Trang et al., 2007) and *(ii)* for acoustic devices for which the resonant frequency or the velocity is dependent on the dimensional values (Lamb wave devices, thickness and shear mode), temperature can affect changes of both material properties and boundary conditions(Kar & Bose, 2011; Taborda et al., 2015).

2.7 Correlation Between Optical Properties and Acoustic Wave Velocity in Mo/AlN SAW Devices

In Figure 19 shows the relationship between optical properties (reflectance), acoustic wave velocity and deposition temperature in AlN films. In Figure 19 it is clearly shown that the reduction of grain size, reduction of reflectance and the increase of the color purity when the deposition temperature decreases, is associated with an increasing of frequency response, due to the increase of acoustic wave velocity, produced by energy activation at the higher deposition temperature. From this correlation, it is possible to determine that one merit index associates to an acceptable reflectance and acoustic wave velocity at the same temperature deposition.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Is observed a direct relationship between the roughnesses of the AlN thin films and increasing the deposition temperature, the AFM results show a decrease in roughness with increasing temperature in the deposit. Figure 19b shows a clear correlation between the roughness and the Rayleigh velocity, obtaining the highest velocity for the samples grown at a higher temperature of 630 °C to deposit are the lower roughness.

Therefore, the AlN films deposited around 450 ºC offer the best synergy for optical and acoustic properties with good reflectance and an acceptable acoustic wave velocity. Furthermore, the nanometric grain size and in thickness of such thin films AlN/Si_3N_4 a low surface roughness (≤120 nm), the nucleation side is characterized by a relatively small grain size (≤70 nm), which can be decreased with increased substrate temperature improving processes nucleation, hence enhancing the propagation of elastic waves as shown in Table 5.

The morphological properties of AlN are highly dependent on their polarity, since the nitrogen, finished films typically have a greater roughness than the finished aluminum, but instead its crystal quality is better. Initial reservoir conditions greatly affect the polarity of the material obtained (Levinshtein et al.,

Figure 19. Correlation between optical properties (Reflectance) and acoustic wave velocity in Mo/AlN SAW devices. In Figure 19b. There is a correlation between roughness and Rayleigh velocity as a function of temperature deposit AlN thin films

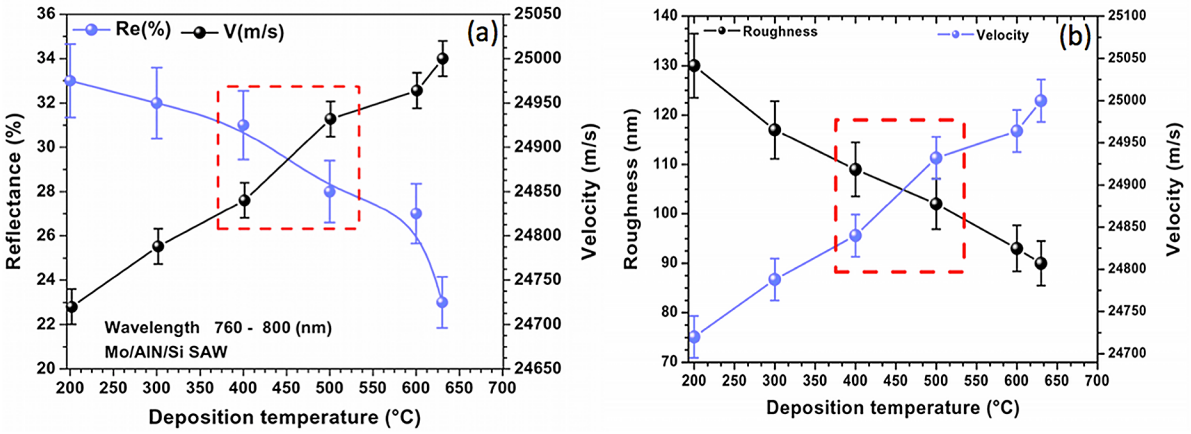


Table 5. Influence of the substrate temperature of AlN thin films on the morphological properties - roughness and the velocity of sound on the SAW

Temperature Substrate (ºC)	Insertion Loss (dB)	Frequency (Ghz)	Grain Size (nm)	Roughness (nm)	SAW velocity (m/s)
200	19,04	0,618	62	130	24720
300	16,45	0,620	56	117	24788
400	12,07	0,621	50	109	24840
500	8,1	0,623	42	102	24932
600	6,09	0,624	38	93	24964
630	5,16	0,625	36	90	25000

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2001). Although in principle it might be thought that all the material has the same polarity, however, is possible that their areas in which the crystals have different polarities. This may explain why the lower the roughness (due to increased substrate temperature during the deposit) on AlN films has favored an increase in the velocity of the waves in the SAW devices.

CONCLUSION

The method of plasma diagnosis by tuning the formation energy of materials formation as HA films may lead to propose and confirm experimental and theoretical models. It was demonstrated that knowledge of the plasma geometry and the corresponding energetic characteristics was a necessary pre-condition for controlling the film stoichiometry and crystallinity adequate for biomedical applications. Furthermore, the knowledge and control of plasma properties is more important than the use of heater and bias used in the conventional sputtering systems.

Likewise, investigated the role of deposition temperature in determining the optical properties of AlN films and the frequency response characteristics of Mo/AlN/Si_3N_4 SAW devices, due to AlN films were grown on Si_3N_4 (100) substrates by PLD using an Al target in N_2 as working gas. The temperature deposition was varied from 200 °C to 630 °C to analyze their effect on the optical and morphological properties of the films. X-ray photoelectron spectroscopy (XPS) confirmed the formation of the binary films. The study revealed that temperature deposition has a marked influence on the chemical and physical nature of all the AlN thin films (oxygen concentration grain size and roughness). For piezoelectric films with a smooth surface, a decrease in the RMS roughness can lead to an increase in its SAW velocity. The evidence is increased with decreasing speed of the roughness and grain size with increasing substrate temperature. It was found a decrease in the reflectance of 10%, an increasing of color purity of 4.5% and dominant wavelength around 1.5% with temperature deposition between 200 - 630 °C. Furthermore, the deposition temperature clearly affects the reduction of micro-drops; this can be possible due to great surface mobility when the temperature was varied from 200 to 630 °C. The SAW propagation velocity is strongly dependent on deposition conditions. The Rayleigh mode and can be obtained with h/λ ratio larger than 0.004. The reflection signal amplitude and phase velocity of Rayleigh wave have been increased with increasing of deposition temperature. Finally, it was possible to design an interdigital structure with Mo circuit on AlN films like an insulator, which can be used as surface acoustic wave (SAW) devices.

Recently authors like Aissa et al.(Aissa et al., 2015) have achieved a breakthrough in obtaining high quality thin films of AlN. The films were deposited by dc magnetron sputtering (dcMS) and high power impulse magnetron sputtering (HiPIMS) deposition techniques. In both cases of films deposited by dcMS and HiPIMS, the XRD results showed that the obtained films are oriented, with full width at half maximum rocking curves of around 1°. Raman spectroscopy revealed higher residual stress relaxation in the AlN epilayers grown by HiPIMS compared to AlN grown by dcMS, highlighted by a blue shift in the E2 (high) Raman mode. The SAW measurements indicated an insertion loss of AlN-SAW devices of about 53 and 35 dB for the AlN films deposited by dcMS and HiPIMS respectively. Even though the values of the insertion loss of the as-deposited AlN thin film by HiPIMS are a little bit inadequate for real applications, these frequency-response characteristics, manifestly, show the potential for high-frequency applications of AlN thin films deposited on Silicon substrates at room temperature.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Aissa, K. A., Achour, A., Elmazria, O., Simon, Q., Elhosni, M., Boulet, P., ... Djouadi, M. (2015). AlN films deposited by dc magnetron sputtering and high power impulse magnetron sputtering for SAW applications. *Journal of Physics. D, Applied Physics*, *48*(14), 145307. doi:10.1088/0022-3727/48/14/145307

Akazawa, H., & Ueno, Y. (2013). Growth of preferentially c-axis oriented hydroxyapatite thin films on Si (100) substrate by electron-cyclotron-resonance plasma sputtering. *Applied Surface Science*, *276*, 217–222. doi:10.1016/j.apsusc.2013.03.070

Akazawa, H., & Ueno, Y. (2014). Control of composition and crystallinity in hydroxyapatite films deposited by electron cyclotron resonance plasma sputtering. *Journal of Physics and Chemistry of Solids*, *75*(1), 94–99. doi:10.1016/j.jpcs.2013.08.014

Astala, R., & Stott, M. (2008). First-principles study of hydroxyapatite surfaces and water adsorption. *Physical Review B: Condensed Matter and Materials Physics*, *78*(7), 075427. doi:10.1103/PhysRevB.78.075427

Bauer, S., Schmuki, P., von der Mark, K., & Park, J. (2013). Engineering biocompatible implant surfaces: Part I: Materials and surfaces. *Progress in Materials Science*, *58*(3), 261–326. doi:10.1016/j.pmatsci.2012.09.001

Belabbes, A., de Carvalho, L., Schleife, A., & Bechstedt, F. (2011). Cubic inclusions in hexagonal AlN, GaN, and InN: Electronic states. *Physical Review B: Condensed Matter and Materials Physics*, *84*(12), 125108. doi:10.1103/PhysRevB.84.125108

Boyd, A., Duffy, H., McCann, R., Cairns, M., & Meenan, B. (2007). The Influence of argon gas pressure on co-sputtered calcium phosphate thin films. *Nuclear Instruments & Methods in Physics Research. Section B, Beam Interactions with Materials and Atoms*, *258*(2), 421–428. doi:10.1016/j.nimb.2007.02.072

Boyd, A., Meenan, B., & Leyland, N. (2006). Surface characterisation of the evolving nature of radio frequency (RF) magnetron sputter deposited calcium phosphate thin films after exposure to physiological solution. *Surface and Coatings Technology*, *200*(20-21), 6002–6013. doi:10.1016/j.surfcoat.2005.09.032

Caicedo, J., Pérez, J., Caicedo, H., & Riascos, H. (2013). Determination of physical response in (Mo/AlN) SAW devices. *Surface Review and Letters*, *20*(2), 1350017. doi:10.1142/S0218625X13500170

Caicedo, J., Zambrano, G., Aperador, W., Escobar-Alarcon, L., & Camps, E. (2011). Mechanical and electrochemical characterization of vanadium nitride (VN) thin films. *Applied Surface Science*, *258*(1), 312–320. doi:10.1016/j.apsusc.2011.08.057

Carradò, A. (2011). Nano-crystalline pulsed laser deposition hydroxyapatite thin films on Ti substrate for biomedical application. *Journal of Coatings Technology and Research*, *8*(6), 749–755. doi:10.100711998-011-9355-9

Chen, F. F., & Smith, M. D. (1984). *Plasma*. Wiley Online Library.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chung, R. J., Hsieh, M. F., Huang, C. W., Perng, L. H., Wen, H. W., & Chin, T. S. (2006). Antimicrobial effects and human gingival biocompatibility of hydroxyapatite sol–gel coatings. *Journal of Biomedical Materials Research. Part B, Applied Biomaterials*, *76*(1), 169–178. doi:10.1002/jbm.b.30365 PMID:16080164

Clement, M., Vergara, L., Sangrador, J., Iborra, E., & Sanz-Hervás, A. (2004). SAW characteristics of AlN films sputtered on silicon substrates. *Ultrasonics,* 42(1–9), 403-407. doi:10.1016/j.ultras.2004.01.034

Cressler, J. D. (2005). *Silicon Heterostructure Handbook: Materials, Fabrication, Devices, Circuits and Applications of SiGe and Si Strained-Layer Epitaxy*. CRC press. doi:10.1201/9781420026580

Dong, Z. L., Khor, K. A., Quek, C., White, T. J., & Cheang, P. (2003). TEM and STEM analysis on heat-treated and in vitro plasma-sprayed hydroxyapatite/Ti-6Al-4V composite coatings. *Biomaterials*, *24*(1), 97–105. doi:10.1016/S0142-9612(02)00267-3 PMID:12417183

Dos Santos, E., Moldovan, M., Jacomine, L., Mateescu, M., Werckmann, J., Anselme, K., ... Pelletier, H. (2010). Oriented hydroxyapatite single crystals produced by the electrodeposition method. *Materials Science and Engineering B*, *169*(1), 138–144. doi:10.1016/j.mseb.2009.10.007

Drafts, B. (2001). Acoustic wave technology sensors. *IEEE Transactions on Microwave Theory and Techniques*, *49*(4), 795–802. doi:10.1109/22.915466

Ellies, L., Nelson, D., & Featherstone, J. (1992). Crystallographic changes in calcium phosphates during plasma-spraying. *Biomaterials*, *13*(5), 313–316. doi:10.1016/0142-9612(92)90055-S PMID:1318087

Feddes, B., Wolke, J., Jansen, J., & Vredenberg, A. (2003). Radio frequency magnetron sputtering deposition of calcium phosphate coatings: The effect of resputtering on the coating composition. *Journal of Applied Physics*, *93*(12), 9503–9507. doi:10.1063/1.1576894

Ferro, D., Barinov, S. M., Rau, J. V., Teghil, R., & Latini, A. (2005). Calcium phosphate and fluorinated calcium phosphate coatings on titanium deposited by Nd: YAG laser at a high fluence. *Biomaterials*, *26*(7), 805–812. doi:10.1016/j.biomaterials.2004.03.027 PMID:15350786

Guipont, V., Jeandin, M., Bansard, S., Khor, K. A., Nivard, M., Berthe, L., ... Boustie, M. (2010). Bond strength determination of hydroxyapatite coatings on Ti-6Al-4V substrates using the LAser Shock Adhesion Test (LASAT). *Journal of Biomedical Materials Research. Part A*, *95*(4), 1096–1104. doi:10.1002/jbm.a.32907 PMID:20878900

Hadjoub, Z., Touati, I., Doghmane, M., & Doghmane, A. (2008). Application of negative velocity dispersion curves to the distinction between layer and substrate Rayleigh waves. *Comptes Rendus Physique*, *9*(8), 903–910. doi:10.1016/j.crhy.2008.08.001

Hahn, B.-D., Lee, J.-M., Park, D.-S., Choi, J.-J., Ryu, J., Yoon, W.-H., ... Kim, H.-E. (2011). Enhanced bioactivity and biocompatibility of nanostructured hydroxyapatite coating by hydrothermal annealing. *Thin Solid Films*, *519*(22), 8085–8090. doi:10.1016/j.tsf.2011.07.008

Han, J. G. (2009). Recent progress in thin film processing by magnetron sputtering with plasma diagnostics. *Journal of Physics. D, Applied Physics*, *42*(4). doi:10.1088/0022-3727/42/4/043001

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hashimoto, K.-Y., & Yamaguchi, M. (2001). Excitation and propagation of shear-horizontal-type surface and bulk acoustic waves. *IEEE Transactions on* Ultrasonics, Ferroelectrics, and Frequency Control, *48*(5), 1181–1188.

Hollister, S. J. (2005). Porous scaffold design for tissue engineering. *Nature Materials*, *4*(7), 518–524. doi:10.1038/nmat1421 PMID:16003400

Hong, Z., Luan, L., Paik, S.-B., Deng, B., Ellis, D. E., Ketterson, J. B., ... Rossi, A. (2007). Crystalline hydroxyapatite thin films produced at room temperature—an opposing radio frequency magnetron sputtering approach. *Thin Solid Films*, *515*(17), 6773–6780. doi:10.1016/j.tsf.2007.02.089

Iliescu, M., Nelea, V., Werckmann, J., & Mihailescu, I. (2004). Transmission electron microscopy investigation of pulsed-laser deposited hydroxylapatite thin films prepared by tripod and focused ion beam techniques. *Surface and Coatings Technology*, *187*(1), 131–140. doi:10.1016/j.surfcoat.2004.01.022

Jallot, E., Nedelec, J.-M., Grimault, A. S., Chassot, E., Grandjean-Laquerriere, A., Laquerriere, P., & Laurent-Maquin, D. (2005). STEM and EDXS characterisation of physico-chemical reactions at the periphery of sol–gel derived Zn-substituted hydroxyapatites during interactions with biological fluids. *Colloids and Surfaces. B, Biointerfaces*, *42*(3), 205–210. doi:10.1016/j.colsurfb.2005.03.001 PMID:15893220

Jedynski, M., Hoffman, J., Mroz, W., & Szymanski, Z. (2008). Plasma plume induced during ArF laser ablation of hydroxyapatite. *Applied Surface Science*, *255*(5), 2230–2236. doi:10.1016/j.apsusc.2008.07.070

Jeong, Y.-H., Choe, H.-C., & Eun, S.-W. (2011). Hydroxyapatite coating on the Ti–35Nb–xZr alloy by electron beam-physical vapor deposition. *Thin Solid Films*, *519*(20), 7050–7056. doi:10.1016/j.tsf.2011.04.086

Jiang, L., Shen, W., Ogawa, H., & Guo, Q. (2003). Temperature dependence of the optical properties in hexagonal AlN. *Journal of Applied Physics*, *94*(9), 5704–5709. doi:10.1063/1.1616988

Joanni, E., Ferro, M. C., Mardare, C. C., Mardare, A. I., Fernandes, J. R. A., & Pina, S. C. A. (2004). Pulsed laser deposition of SiO2-P2O5-CaO-MgO glass coatings on titanium substrates. *Materials Research*, *7*(3), 431–436. doi:10.1590/S1516-14392004000300010

Jung, J.-P., Lee, J.-B., Kim, J.-S., & Park, J.-S. (2004). Fabrication and characterization of high frequency SAW device with IDT/ZnO/AlN/Si conFigureuration: Role of AlN buffer. *Thin Solid Films*, *447*, 605–609. doi:10.1016/j.tsf.2003.07.022

Kanzaki, N., Treboux, G., Onuma, K., Tsutsumi, S., & Ito, A. (2001). Calcium phosphate clusters. *Biomaterials*, *22*(21), 2921–2929. doi:10.1016/S0142-9612(01)00039-4 PMID:11561898

Kao, K.-S., Cheng, C.-C., & Chen, Y.-C. (2002). Synthesis and surface acoustic wave properties of AlN films deposited on LiNbO/sub 3/substrates. *Ultrasonics, Ferroelectrics, and Frequency Control. IEEE Transactions on*, *49*(3), 345–349.

Kar, J. P., & Bose, G. (2011). *Aluminum Nitride (AlN) film based acoustic devices: material synthesis and device fabrication*. INTECH Open Access Publisher.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Khor, K., Dong, Z., Quek, C., & Cheang, P. (2000). Microstructure investigation of plasma sprayed HA/Ti6Al4V composites by TEM. *Materials Science and Engineering A*, *281*(1), 221–228. doi:10.1016/S0921-5093(99)00717-0

Kim, W.-G., & Choe, H.-C. (2011). Surface characteristics of hydroxyapatite/titanium composite layer on the Ti-35Ta-xZr surface by RF and DC sputtering. *Thin Solid Films*, *519*(20), 7045–7049. doi:10.1016/j.tsf.2011.04.090

Layrolle, P., Ito, A., & Tateishi, T. (1998). Sol-Gel Synthesis of Amorphous Calcium Phosphate and Sintering into Microporous Hydroxyapatite Bioceramics. *Journal of the American Ceramic Society*, *81*(6), 1421–1428. doi:10.1111/j.1151-2916.1998.tb02499.x

Levinshtein, M. E., Rumyantsev, S. L., & Shur, M. S. (2001). *Properties of Advanced Semiconductor Materials: GaN, AIN, InN, BN, SiC, SiGe*. John Wiley & Sons.

Li, C., Hong, G., Yu, H., & Qi, L. (2010). Facile fabrication of honeycomb-patterned thin films of amorphous calcium carbonate and mosaic calcite. *Chemistry of Materials*, *22*(10), 3206–3211. doi:10.1021/cm100363a

Li, C., & Qi, L. (2008). Bioinspired fabrication of 3D ordered macroporous single crystals of calcite from a transient amorphous phase. *Angewandte Chemie International Edition*, *47*(13), 2388–2393. doi:10.1002/anie.200705403 PMID:18300213

Lieberman, M., & Lichtenberg, A. (2005). *Principles of Plasma Discharges and Plasma Processing*. John Wiley and Sons.

Lieberman, M. A., & Lichtenberg, A. J. (1994). Principles of plasma discharges and materials processing. *MRS Bulletin*, *30*, 899–901.

Liu, D., Savino, K., & Yates, M. Z. (2011). Coating of hydroxyapatite films on metal substrates by seeded hydrothermal deposition. *Surface and Coatings Technology*, *205*(16), 3975–3986. doi:10.1016/j.surfcoat.2011.02.008

López, E. O., Mello, A., Sendão, H., Costa, L. T., Rossi, A. L., Ospina, R. O., ... Rossi, A. M. (2013). Growth of crystalline hydroxyapatite thin films at room temperature by tuning the energy of the RF-magnetron sputtering plasma. *ACS Applied Materials & Interfaces*, *5*(19), 9435–9445. doi:10.1021/am4020007 PMID:24059686

López, E. O., Rossi, A. L., Archanjo, B. S., Ospina, R. O., Mello, A., & Rossi, A. M. (2015). Crystalline nano-coatings of fluorine-substituted hydroxyapatite produced by magnetron sputtering with high plasma confinement. *Surface and Coatings Technology*, *264*, 163–174.

Louis-Achille, V., De Windt, L., & Defranceschi, M. (2000). Electronic structure of minerals: The apatite group as a relevant example. *International Journal of Quantum Chemistry*, *77*(6), 991–1006. doi:10.1002/(SICI)1097-461X(2000)77:6<991::AID-QUA7>3.0.CO;2-Z

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Macaskie, L. E., Yong, P., Paterson-Beedle, M., Thackray, A. C., Marquis, P. M., Sammons, R. L., ... Hall, L. D. (2005). A novel non line-of-sight method for coating hydroxyapatite onto the surfaces of support materials by biomineralization. *Journal of Biotechnology*, *118*(2), 187–200. doi:10.1016/j.jbiotec.2005.03.006 PMID:15964651

Mahmood, A., Machorro, R., Muhl, S., Heiras, J., Castillon, F., Farıas, M., & Andrade, E. (2003). Optical and surface analysis of DC-reactive sputtered AlN films. *Diamond and Related Materials*, *12*(8), 1315–1321. doi:10.1016/S0925-9635(03)00076-1

Maruska, H. P., & Tietjen, J. (1969). The preparation and properties of Vapor-Deposited single-crystalline GaN. *Applied Physics Letters*, *15*(10), 327–329. doi:10.1063/1.1652845

Melde, B. J., & Stein, A. (2002). Periodic macroporous hydroxyapatite-containing calcium phosphates. *Chemistry of Materials*, *14*(8), 3326–3331. doi:10.1021/cm020259b

Mello, A., Hong, Z., Rossi, A., Luan, L., Farina, M., Querido, W., ... Webster, T. (2007). Osteoblast proliferation on hydroxyapatite thin coatings produced by right angle magnetron sputtering. *Biomedical Materials (Bristol, England)*, *2*(2), 67–77. doi:10.1088/1748-6041/2/2/003 PMID:18458438

Miao, S., Cheng, K., Weng, W., Du, P., Shen, G., Han, G., ... Zhang, S. (2008). Fabrication and evaluation of Zn containing fluoridated hydroxyapatite layer with Zn release ability. *Acta Biomaterialia*, *4*(2), 441–446. doi:10.1016/j.actbio.2007.08.013 PMID:17942380

Montazeri, L., Javadpour, J., Shokrgozar, M., Bonakdar, S., Moghaddam, M. K., & Asgary, V. (2013). The interaction of plasma proteins with nano-size fluoride-substituted apatite powders. *Ceramics International*, *39*(6), 6145–6152. doi:10.1016/j.ceramint.2013.01.033

Morgan, H., Wilson, R., Elliott, J., Dowker, S., & Anderson, P. (2000). Preparation and characterisation of monoclinic hydroxyapatite and its precipitated carbonate apatite intermediate. *Biomaterials*, *21*(6), 617–627. doi:10.1016/S0142-9612(99)00225-2 PMID:10701462

Morkoc, H. (2013). *Nitride semiconductors and devices* (Vol. 32). Springer Science & Business Media. doi:10.1002/9783527649006

Mróz, W., Jedyński, M., Prokopiuk, A., Ślósarczyk, A., & Paszkiewicz, Z. (2009). Characterization of calcium phosphate coatings doped with Mg, deposited by pulsed laser deposition technique using ArF excimer laser. *Micron (Oxford, England)*, *40*(1), 140–142. doi:10.1016/j.micron.2008.01.016 PMID:18407507

Narayanan, R., Seshadri, S., Kwon, T., & Kim, K. (2008). Calcium phosphate-based coatings on titanium and its alloys. *Journal of Biomedical Materials Research. Part B, Applied Biomaterials*, *85*(1), 279–299. doi:10.1002/jbm.b.30932 PMID:17853421

Nelea, V., Morosanu, C., Iliescu, M., & Mihailescu, I. (2003). Microstructure and mechanical properties of hydroxyapatite thin films grown by RF magnetron sputtering. *Surface and Coatings Technology*, *173*(2), 315–322. doi:10.1016/S0257-8972(03)00729-1

Ozeki, K., Aoki, H., & Masuzawa, T. (2010). TEM Observation of Nano-Interface between Hydroxyapatite Sputtered Film and Titanium Substrate. *Journal of the Australasian Ceramic Society*, *46*(2), 59–62.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Padilla, S., Roman, J., Carenas, A., & Vallet-Regı, M. (2005). The influence of the phosphorus content on the bioactivity of sol–gel glass ceramics. *Biomaterials*, *26*(5), 475–483. doi:10.1016/j.biomaterials.2004.02.054 PMID:15276355

Paduano, Q., Weyburne, D., Jasinski, J., & Liliental-Weber, Z. (2004). Effect of initial process conditions on the structural properties of AlN films. *Journal of Crystal Growth*, *261*(2), 259–265. doi:10.1016/j.jcrysgro.2003.11.017

Paul, W., & Sharma, C. P. (2006). Nanoceramic matrices: Biomedical applications. *Am J Biochem Biotechnol*, *2*(2), 41–48. doi:10.3844/ajbbsp.2006.41.48

Pelletier, H., Carradò, A., Faerber, J., & Mihailescu, I. N. (2011). Microstructure and mechanical characteristics of hydroxyapatite coatings on Ti/TiN/Si substrates synthesized by pulsed laser deposition. *Applied Physics. A, Materials Science & Processing*, *102*(3), 629–640. doi:10.100700339-010-6144-8

Pereiro, I., Rodríguez-Valencia, C., Serra, C., Solla, E., Serra, J., & González, P. (2012). Pulsed laser deposition of strontium-substituted hydroxyapatite coatings. *Applied Surface Science*, *258*(23), 9192–9197. doi:10.1016/j.apsusc.2012.04.063

Pérez, J., Riascos, H., Caicedo, J., Cabrera, G., & Yate, L. (2011a). *Optical Properties Dependence with Gas Pressure in AlN Films Deposited by Pulsed Laser Ablation. Paper presented at the Journal of Physics: Conference Series*. 10.1088/1742-6596/274/1/012119

Pérez, J., Riascos, H., Caicedo, J., Cabrera, G., & Yate, L. (2011b). Optical Properties Dependence with Gas Pressure in AlN Films Deposited by Pulsed Laser Ablation. *Journal of Physics: Conference Series*, *274*(1), 012119. doi:10.1088/1742-6596/274/1/012119

Perez Taborda, J. A., Riascos, H., & Vera, L. (n. d.). Correlation between optical, morphological and compositional properties of Aluminum Nitride thin films by Pulsed Laser Deposition.

Pérez Taborda, J. A., Riascos Landázuri, H., Jiménez García, F. N., & Caicedo Angulo, J. C. (2010). Structural and morphological analysis of aluminum nitride films of obtained by pulsed laser deposition. *Ciencia e Ingeniería Neogranadina*, *20*(2), 107–115.

Rabiei, A., Thomas, B., Jin, C., Narayan, R., Cuomo, J., Yang, Y., & Ong, J. (2006). A study on functionally graded HA coatings processed using ion beam assisted deposition with in situ heat treatment. *Surface and Coatings Technology*, *200*(20), 6111–6116. doi:10.1016/j.surfcoat.2005.09.027

Raole, P., Mukherjee, S., & John, P. (2005). X-ray photoelectron spectroscopic study of plasma source nitrogen ion implantation in single crystal natural diamond. *Diamond and Related Materials*, *14*(3), 482–485. doi:10.1016/j.diamond.2004.11.009

Rau, J., Smirnov, V., Laureti, S., Generosi, A., Varvaro, G., Fosca, M., ... Barinov, S. (2010). Properties of pulsed laser deposited fluorinated hydroxyapatite films on titanium. *Materials Research Bulletin*, *45*(9), 1304–1310. doi:10.1016/j.materresbull.2010.04.028

Rosenberger, L., Baird, R., McCullen, E., Auner, G., & Shreve, G. (2008). XPS analysis of aluminum nitride films deposited by plasma source molecular beam epitaxy. *Surface and Interface Analysis*, *40*(9), 1254–1261. doi:10.1002ia.2874

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ruterana, P., Albrecht, M., & Neugebauer, J. (2006). *Nitride semiconductors: handbook on materials and devices*. John Wiley & Sons.

Saju, K., Reshmi, R., Jayadas, N., James, J., & Jayaraj, M. (2009). Polycrystalline coating of hydroxyapatite on TiAl6V4 implant material grown at lower substrate temperatures by hydrothermal annealing after pulsed laser deposition. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine*, *223*(8), 1049–1057. doi:10.1243/09544119JEIM568 PMID:20092101

Schwertfeger, F., Emmerling, A., Gross, J., Schubert, U., Fricke, J., & Attia, Y. (1994). *Sol-Gel Processing and Applications* (p. 343). New York. doi:10.1007/978-1-4615-2570-7_29

Shaginyan, L. R., Han, J. G., Shaginyan, V. R., & Musil, J. (2006). Evolution of film temperature during magnetron sputtering. *Journal of Vacuum Science & Technology A, 24*(4), 1083-1090. doi:10.1116/1.2210947

Shi, J., Chen, C., Yu, H., & Zhang, S. (2008). Application of magnetron sputtering for producing bioactive ceramic coatings on implant materials. *Bulletin of Materials Science*, *31*(6), 877–884. doi:10.100712034-008-0140-z

Shun'ko, E. V. (2009). *Langmuir probe in theory and practice*. Universal-Publishers.

Snyders, R., Bousser, E., Music, D., Jensen, J., Hocquet, S., & Schneider, J. M. (2008). Influence of the Chemical Composition on the Phase Constitution and the Elastic Properties of RF-Sputtered Hydroxyapatite Coatings. *Plasma Processes and Polymers*, *5*(2), 168–174. doi:10.1002/ppap.200700112

Snyders, R., Music, D., Sigumonrong, D., Schelnberger, B., Jensen, J., & Schneider, J. (2007). Experimental and ab initio study of the mechanical properties of hydroxyapatite. *Applied Physics Letters*, *90*(19), 193902. doi:10.1063/1.2738386

Socol, G., Torricelli, P., Bracci, B., Iliescu, M., Miroiu, F., Bigi, A., ... Mihailescu, I. (2004). Biocompatible nanocrystalline octacalcium phosphate thin films obtained by pulsed laser deposition. *Biomaterials*, *25*(13), 2539–2545. doi:10.1016/j.biomaterials.2003.09.044 PMID:14751739

Strite, S., & Morkoç, H. (1992). GaN, AlN, and InN: A review. *Journal of Vacuum Science & Technology B Microelectronics and Nanometer Structures*, *10*(4), 1237–1266. doi:10.1116/1.585897

Surmenev, R. A. (2012). A review of plasma-assisted methods for calcium phosphate-based coatings fabrication. *Surface and Coatings Technology*, *206*(8), 2035–2056. doi:10.1016/j.surfcoat.2011.11.002

Surmenev, R. A., Surmeneva, M. A., Evdokimov, K. E., Pichugin, V. F., Peitsch, T., & Epple, M. (2011). The influence of the deposition parameters on the properties of an rf-magnetron-deposited nanostructured calcium phosphate coating and a possible growth mechanism. *Surface and Coatings Technology*, *205*(12), 3600–3606. doi:10.1016/j.surfcoat.2010.12.039

Surmeneva, M. A., Chaikina, M. V., Zaikovskiy, V. I., Pichugin, V. F., Buck, V., Prymak, O., ... Surmenev, R. A. (2013). The structure of an RF-magnetron sputter-deposited silicate-containing hydroxyapatite-based coating investigated by high-resolution techniques. *Surface and Coatings Technology*, *218*, 39–46. doi:10.1016/j.surfcoat.2012.12.023

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Taborda, J. A. P., Caicedo, J., Grisales, M., Saldarriaga, W., & Riascos, H. (2015). Deposition pressure effect on chemical, morphological and optical properties of binary Al-nitrides. *Optics & Laser Technology*, *69*, 92–103. doi:10.1016/j.optlastec.2014.12.009

Touati, I., Hadjoub, Z., & Doghmane, A. (2006). Investigations of elastic properties of Al2O3 layers on sic substrate via acoustic microscopie. *Physical & chemical news*, 27, 26-30.

Trang, H., Marie-Helene, V., Jean-Sebastien, D., Philippe, R., Philippe, B., & Pierre, L.-A. (2007). Temperature-compensated structure for SAW pressure sensor in very high temperature. *Paper presented at the Frequency Control Symposium, 2007 Joint with the 21st European Frequency and Time Forum*. IEEE International. 10.1109/FREQ.2007.4319026

Treboux, G., Layrolle, P., Kanzaki, N., Onuma, K., & Ito, A. (2000). Existence of Posner's cluster in vacuum. *The Journal of Physical Chemistry A*, *104*(21), 5111–5114. doi:10.1021/jp994399t

Tredwin, C. J., Young, A. M., Neel, E. A. A., Georgiou, G., & Knowles, J. C. (2014). Hydroxyapatite, fluor-hydroxyapatite and fluorapatite produced via the sol–gel method: Dissolution behaviour and biological properties after crystallisation. *Journal of Materials Science. Materials in Medicine*, *25*(1), 47–53. doi:10.100710856-013-5050-y PMID:24052344

Uchida, M., Oyane, A., Kim, H. M., Kokubo, T., & Ito, A. (2004). Biomimetic Coating of Laminin–Apatite Composite on Titanium Metal and Its Excellent Cell-Adhesive Properties. *Advanced Materials*, *16*(13), 1071–1074. doi:10.1002/adma.200400152

Ueda, K., Narushima, T., Goto, T., Taira, M., & Katsube, T. (2007). Fabrication of calcium phosphate films for coating on titanium substrates heated up to 773 K by RF magnetron sputtering and their evaluations. *Biomedical Materials*, *2*(3), S160–S166. doi:10.1088/1748-6041/2/3/S14 PMID:18458462

Vallet-Regí, M., & Ruiz-Hernández, E. (2011). Bioceramics: From bone regeneration to cancer nanomedicine. *Advanced Materials*, *23*(44), 5177–5218. doi:10.1002/adma.201101586 PMID:22009627

Van Der Wal, E., Oldenburg, S., Heij, T., Van Der Gon, A. D., Brongersma, H., Wolke, J., ... Vredenberg, A. (2006). Adsorption and desorption of Ca and PO 4 species from SBFs on RF-sputtered calcium phosphate thin films. *Applied Surface Science*, *252*(10), 3843–3854. doi:10.1016/j.apsusc.2005.06.006

Van der Wal, E., Wolke, J., Jansen, J., & Vredenberg, A. (2005). Initial reactivity of rf magnetron sputtered calcium phosphate thin films in simulated body fluids. *Applied Surface Science*, *246*(1), 183–192.

Wang, H., Eliaz, N., & Hobbs, L. W. (2011). The nanostructure of an electrochemically deposited hydroxyapatite coating. *Materials Letters*, *65*(15), 2455–2457. doi:10.1016/j.matlet.2011.05.016

Wang, H., Eliaz, N., Xiang, Z., Hsu, H.-P., Spector, M., & Hobbs, L. W. (2006). Early bone apposition in vivo on plasma-sprayed and electrochemically deposited hydroxyapatite coatings on titanium alloy. *Biomaterials*, *27*(23), 4192–4203. doi:10.1016/j.biomaterials.2006.03.034 PMID:16618502

Wen, C., Xu, W., Hu, W., & Hodgson, P. (2007). Hydroxyapatite/titania sol–gel coatings on titanium–zirconium alloy for biomedical applications. *Acta Biomaterialia*, *3*(3), 403–410. doi:10.1016/j.actbio.2006.10.004 PMID:17204459

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Xu, S., Long, J., Sim, L., Diong, C. H., & Ostrikov, K. K. (2005). RF plasma sputtering deposition of hydroxyapatite bioceramics: Synthesis, performance, and biocompatibility. *Plasma Processes and Polymers*, *2*(5), 373–390. doi:10.1002/ppap.200400094

Yamaguchi, T., Tanaka, Y., & Ide-Ektessabi, A. (2006). Fabrication of hydroxyapatite thin films for biomedical applications using RF magnetron sputtering. *Nuclear Instruments & Methods in Physics Research. Section B, Beam Interactions with Materials and Atoms*, *249*(1), 723–725. doi:10.1016/j.nimb.2006.03.126

Yamashita, H., Fukui, K., Misawa, S., & Yoshida, S. (1979). Optical properties of AlN epitaxial thin films in the vacuum ultraviolet region. *Journal of Applied Physics*, *50*(2), 896–898. doi:10.1063/1.326007

Yonggang, Y., Wolke, J. G., Yubao, L., & Jansen, J. A. (2007). In vitro evaluation of different heat-treated radio frequency magnetron sputtered calcium phosphate coatings. *Clinical Oral Implants Research*, *18*(3), 345–353. doi:10.1111/j.1600-0501.2006.01332.x PMID:17298487

Zayats, N., Boĭko, V., Gentsar, P., Litvin, O., Papusha, V., & Sopinskiĭ, N. (2008). Optical studies of AlN/n-Si (100) films obtained by the method of high-frequency magnetron sputtering. *Semiconductors*, *42*(2), 195–198. doi:10.1134/S1063782608020139

Zhang, K., Francis, L. F., Yan, H., & Stein, A. (2005). Apatite Converted from 3-D Ordered Macroporous Sol–Gel Bioactive Glass (3DOM-BG) Particles. *Journal of the American Ceramic Society*, *88*(3), 587–592. doi:10.1111/j.1551-2916.2005.00143.x

Zhang, S. (2013). *Hydroxyapatite Coatings for Biomedical Applications*. CRC Press.

Zhang, S., Wang, Y., Zeng, X., Cheng, K., Qian, M., Sun, D., ... Chia, W. (2007). Evaluation of interfacial shear strength and residual stress of sol–gel derived fluoridated hydroxyapatite coatings on Ti6Al4V substrates. *Engineering Fracture Mechanics*, *74*(12), 1884–1893. doi:10.1016/j.engfracmech.2006.04.021

Zhuang, C., Zhao, J., Jia, F., Guan, C., Wu, Z., Bai, Y., & Jiang, X. (2009). Tuning bond contents in B–C–N films via temperature and bias voltage within RF magnetron sputtering. *Surface and Coatings Technology*, *204*(5), 713–717. doi:10.1016/j.surfcoat.2009.09.031

This research was previously published in Research Perspectives on Functional Micro- and Nanoscale Coatings edited by Ana Zuzuarregui and Maria Carmen Morant-Miñana; pages 136-169; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 57
Optimal Intervention Methods for Markovian Gene Regulatory Networks

Mohammadmahdi Rezaei Yousefi
The Ohio State University, USA

ABSTRACT

A central problem in translational medicine is to provide a framework for deriving and studying effective intervention methods to elicit desired steady-state behavior for a gene regulatory network of interest with Markovian dynamics. Heretofore, two rather different external control approaches have been taken. The first optimizes a subjectively defined cost function while modeling treatment constraints; therefore, desirable shift of the steady-state mass is a by-product. The second approach, on the other hand, focuses solely on the steady-state behavior of the network and provides the maximal shift achievable. Although both approaches are optimal with respect to their objectives, the choice of which to use depends on the treatment goals.

INTRODUCTION

In order for a cancer treatment to be successful one must account for the patient's condition, the nature of the disease, its dynamics, the effect of antitumor drugs or radiation, the duration of effect and side effects. Each of these is a major field in the medical community, and a vast amount effort has been devoted to their characterization. Typically, the main objective of many treatment schemes is to find the right dose schedules while maximizing the benefit to toxicity ratio for a large portion of the patient population (Simon & Norton, 2006). To this end, treatments usually involve some kind of protocol and several drug candidates acting on various gene products with the aim of mitigating undesirable gene functions in cancer-related genetic pathways. Cancer therapies, for example chemotherapy, are usually administered in cycles, where each cycle begins by delivering a drug that kills cells which divide rapidly. The drug is effective for some period, on both normal and tumor cells, after which there is usually a rest period

DOI: 10.4018/978-1-7998-1204-3.ch057

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

allowing the body to recover from the adverse effects of the drug (Brunton, Chabner, & Knollmann, 2010). To maximize the chance of eradicating a tumor, one should deliver the most effective drug at a sufficient dosage over as short a time as possible (Simon & Norton, 2006). However, since the toxicity of a drug could be intolerable to the patient, we desire a therapeutic method that acts very rapidly and with high efficiency so that a large percentage of tumor cells die or shift into a mode where they can no longer proliferate. The method should, at the same time, prolong patients' lives with reasonable quality during and after treatment.

The duration of effective response for a given set of drugs depends on many factors that may vary among a patient population or even in a population of cells within an individual. Achieving an acceptable success rate is tied to three key factors: (1) building accurate models for the underlying dynamical system governing the cancer and the drug response; (2) forcing cancerous cells to either initiate cell death or terminate differentiation as early as possible; and (3) choosing a suitable intervention strategy according to the states of the tumor and patient at the beginning of each therapeutic cycle (Yousefi, Datta, & Dougherty, 2013).

A gene regulatory network (GRN) refers to a class of models describing the multivariate functional relationships among a cohort of genes or their products. From a graphical perspective, genes are nodes in this network and edges describe regulatory relationships between genes. A GRN can be alternatively viewed as an input-output system, where the inputs are signals from signaling pathways or transcription factors, and the outputs are gene expression levels at a given time (Huang, Tienda-Luna, & Wang, 2009). A regulatory network should be able to predict the behavior of the collection of genes of interest under different conditions and perturbations (Dougherty & Shmulevich, 2012; Pe'er & Hacohen, 2011). These networks aim to model cellular control and how abnormal cell functions result from one or several breakdowns in the regulatory mechanisms, making them an important part of translational medicine (Dougherty, Pal, Qian, Bittner, & Datta, 2010).

Building a model for the underlying GRN is the first step toward designing optimal intervention. There has been a considerable amount of work on developing models and inference methods that provide accuracy, robustness, adaptivity and a system-level view, including Bayesian networks (Friedman, 2004), signed directed graphs (Kyoda, Baba, Onami, & Kitano, 2004), Petri nets (Chaouiya, Remy, & Thieffry, 2008; Matsuno, Doi, Nagasaki, & Miyano, 2000), differential equations (De Jong, 2002), stochastic differential equations (Chen, Wang, Tseng, Huang, & Kao, 2005; El Samad, Khammash, Petzold, & Gillespie, 2005), Boolean networks (Kauffman, 1969) and probabilistic Boolean networks (PBNs) (Shmulevich, Dougherty, Kim, & Zhang, 2002). Each modeling paradigm has its own strengths and weaknesses. For example, Bayesian networks are descriptive and require a large number of samples for inference. Boolean models are predictive, but they are very simplistic and come short when a more refined characterization of a regulatory process is needed. Although ordinary differential equations can capture quantitative molecular or physical details of the system, predictions are usually accurate only when the number of molecular species involved is large. When regulatory molecules are produced in small concentrations, fluctuations in the number of molecules have a significant impact on the process dynamics, and therefore a stochastic differential equation model is more appropriate (Cao & Samuels, 2009; Shmulevich & Aitchison, 2009).

Modeling GRNs via Markovian dynamical networks, such as PBNs, has received much attention recently because they are simple and can capture uncertainty, intrinsic or extrinsic, due to measurement noise in the data, latent variables or stochastic interactions among genes or gene products at different levels. Another feature of PBNs, which makes them attractive for studying intervention methods that affect

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the dynamics of a GRN, is that they can be represented as finite-state Markov chains. Each state in the resulting Markov chain represents a *gene activity profile* (GAP) and state transitions are made according to a *transition probability matrix* (TPM). To be self-contained, this chapter will give a review of PBNs.

The set of all GAPs constitutes the state space of the network, which can be then partitioned into two disjoint sets of *desirable* and *undesirable* cellular states according to the expression levels of a given set of genes. An undesirable state can be associated with a phenotype representative of a cancerous state; for example, cell growth in the absence of growth factors is undesirable (Faryabi, Vahedi, Chamberland, Datta, & Dougherty, 2008). Once a model for the underlying GRN is assumed and desirable and undesirable states are identified, the goal in optimal intervention design is to find therapeutic strategies (control policies) associated with some predefined objective functions such that the evolution of the network is optimally steered from undesirable states to desirable ones (Datta & Dougherty, 2006; Dougherty et al., 2010; Dougherty & Datta, 2005). This problem has been extensively studied within two basic intervention approaches in the context of PBNs, namely structural intervention and external control.

Structural intervention, in its simplest form, involves a one-time change in the regulatory pathway of interest. Given a cancer-related network and a collection of potential structural interventions, the objective is to single out the one resulting in a maximally desirable alteration in the steady-state of the network (Qian & Dougherty, 2008; Shmulevich, Dougherty, & Zhang, 2002; Xiao & Dougherty, 2007).

External control, on the other hand, is generally based on changing the expression status of a set of specific genes or gene products, called the *control* genes, at each decision epoch. The first proposed intervention approach in the context of PBNs was to determine an optimal single-gene perturbation to the network based on first-passage times in Markov chains (Shmulevich, Dougherty, & Zhang, 2002). Attention, later, turned to dynamic-programming-based finite-horizon (Datta, Choudhary, Bittner, & Dougherty, 2003) and infinite-horizon external control methods (Pal, Datta, & Dougherty, 2006). This chapter will be solely focused on external control methods, and will utilize the rich theory of stochastic optimal control.

Decisions regarding possible interventions are usually assumed to be made at every state transition of the network (Datta et al., 2003; Pal et al., 2006). To model dosing constraint, a cyclic approach to intervention may be implemented in which there is a sequence of treatment windows and treatment is allowed only at the beginning of each window (Vahedi, Faryabi, Chamberland, Datta, & Dougherty, 2009). These protocols ignore an important practical aspect of therapy: Due to biological and microenvironmental variability within tumor cells, the action period of an administered drug can vary among a population of patients, so that treatments may have an unknown duration of action. Accounting for this practical modeling assumption introduces an acyclic structure in the intervention protocol. Furthermore, in practice, a physician might monitor a patient at every stage of treatment and decide whether to intervene or not, and if the decision is to intervene, then the patient's body will be under the influence of the drug for some duration, usually followed by a recovery phase. Based on a recently proposed model for tumor growth inhibition, we will first characterize the variability of the length of drug action. Then, we will present a unified approach to both cyclic and acyclic intervention models for PBNs. We accomplish this by placing the problem in the general framework Markov decision processes with partially controlled decision intervals (White, 1978), also known as *semi-Markov* models (Bertsekas, 2005), with an infinite horizon discounting cost objective function. We present a methodology to devise optimal solutions for this class of sequential-decision-making problems and locate optimal intervention policies for synthetically generated gene regulatory networks, and a mutated mammalian cell-cycle network.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Usually, optimal intervention methods are defined relative to a cost value that is assigned through a subjective process. Thus the long-run effect of such methods on a PBN becomes a by-product of the optimization procedure. Since a patient's phenotype is associated with the steady-state behavior of their underlying gene regulatory network, from a practical perspective, it is advantageous to define optimality directly in terms of long-run behavior, which in turn would relieve the physician or biologist from the burden of having to define a subjective cost function (Dougherty et al., 2010; Yousefi & Dougherty, 2013). To address this, we will formulate optimization on the steady-state distribution of PBNs to maximally alter undesirable phenotypes. In addition to desirable and undesirable states, there may exist some "ambiguous" states that are either unrelated to the condition of interest or not fully understood, and thus bear some perceived risk. The framework is flexible enough to account for such cases, where we put constraints on the amount of mass we are willing to accept at different types of states. We will demonstrate the performance of optimal policies using synthetically generated networks as well as a real mammalian cell cycle network and a metastatic melanoma network.

There are other approaches for the control of biological systems. For example, an integrated approach to modeling and control of biological systems has been presented in (Julius et al., 2008), where the authors propose a stochastic hybrid model of the lactose regulation system of *E. coli* bacteria. The dynamics of this model can be further abstracted into a much simpler two-state continuous-time Markov chain, which then can be used to design a feedback controller that drives a certain amount of bacteria to an induced state by manipulating some chemical concentration in the environment (Julius et al., 2008). A recent approach to controlling GRN dynamics has been proposed in (Zañudo & Albert, 2015), where an asynchronous updating scheme for a Boolean network is considered as the model. By integrating the structural and functional information available in the network, the authors propose a procedure to find control targets that directly stabilize the expression or activity of a few select components so that the cell dynamics are driven towards a desired fate or away from an undesired fate. The utility of this approach is demonstrated on a leukemia signaling network and a network controlling the differentiation of T helper cells (Zañudo & Albert, 2015). We refer the interested reader to (Iglesias & Ingalls, 2010; Queinnec, Garcia, Niculescu, & Tarbouriech, 2007) for more examples on the application of control theory in biology.

MARKOVIAN GENE REGULATORY NETWORKS

A "good" model for a GRN should be able to predict the behavior of the main processes, pertinent to the phenotypes of interest, that occur inside a cell or within a population of cells under different conditions and perturbations (Dougherty & Shmulevich, 2012; Pe'er & Hacohen, 2011). It should also identify cellular control mechanisms that lead to normal and abnormal cell functions. Although one can construct complex regulatory models that describe genetic interactions at a fine grain level, this would require a precise understanding of the underlying biological processes and extensive amounts of experimental data to infer the model. Due to these limitations, we may need to move toward coarser models that incorporate uncertainty, rather than complexity, at different levels.

PBNs, as a class of Markovian networks, model uncertainty in gene activity in a GRN (Shmulevich, Dougherty, Kim, et al., 2002). They are essentially a probabilistic extension of a class of rule-based models for the GRNs known as the Boolean networks (Kauffman, 1969). We will first give a brief re-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

view of Boolean networks and then describe a class of PBNs we will be using throughout this chapter. We will also discuss how PBNs can incorporate external intervention inputs, and address two types of controlled PBNs that are central in our study of network intervention.

Boolean Networks

A Boolean network assumes a regulatory graph representing multivariate interactions among a set of genes or gene products. An interaction is specified by a directed edge starting from a predictor gene and ending at an influenced gene. Every gene that is directly connected to another gene may serve as its predictor. A Boolean regulatory rule defines how the predictors of a gene determine its expression level. We restrict ourselves to binary levels, where it is assumed that the expression of each gene is quantized to two levels: zero or one. Zero corresponds to an unexpressed (OFF) gene and one corresponds to an expressed (ON) gene. The mathematical theory extends directly to networks with discrete-valued nodes. We now formally define a Boolean network.

Definition 1: A Boolean network is fully characterized by a set of n nodes, $V=\{v^1,v^2,\ldots,v^n\}$, (representing genes $\{g^1,g^2,\ldots,g^n\}$ or their products) and a collection of Boolean functions $F=\{f_1,f_2,\ldots,f_n\}$ describing the functional relationships between the nodes. Each Boolean function f_i: $\{0,1\}^j \rightarrow \{0,1\}$ determines the value of node i at time $t+1$ given the value of its predictors at time t by

$$v_{t+1}^i = f_i(v_t^{i1}, v_t^{i2}, \ldots, v_t^{ij_i}), \tag{1}$$

where $\{v_t^{i1}, v_t^{i2}, \ldots, v_t^{ij_i}\}$ corresponds to values of j_i predictors of gene g^i, $\{g^{i1},g^{i2},\ldots,g^{ij}\}$, at time t. In a Boolean network, all genes are assumed to update synchronously, so that the network evolves according to a genome-wide GAP, $\mathbf{v}_t = (v_t^1, v_t^2, \ldots, v_t^n)$ at each time t, forming a *state space* of size 2^n. One may alternatively convert each $\mathbf{v}_t$ to a decimal value by $x_t = \sum_{i=1}^{n} 2^{n-i} v_t^i$, where $x \in \mathcal{S} = \{0,1,\ldots,2^n-1\}$ is a *state*. Given an initial state, a Boolean network will eventually enter and stay in a set of states called an *attractor cycle*. The set of states leading to an attractor cycle is known as a *basin of attraction (BOA)* (Dougherty et al., 2010). Once we have a state space, we can identify phenotypes as the attractors of a Boolean network. We label states as being *desirable* or *undesirable* depending on the type and severity of the phenotype they are associated with. The set of desirable and undesirable states are denoted by $\mathcal{D}$ and $\mathcal{U}$, respectively.

Probabilistic Boolean Networks

Due to latent variables, unknown interactions among genes and measurements noise, one may introduce uncertainty into a Boolean network by allowing a collection of several contextual regulatory functions, instead of a single F, as well as random perturbation of the nodes at every time epoch (Shmulevich, Dougherty, Kim, et al., 2002). The resulting model is a PBN formally defined below.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Definition 2: A probabilistic Boolean network is characterized by a set of n nodes, $V=\{v^1,v^2,\ldots,v^n\}$, and a set of m constituent Boolean networks, $\mathbf{F}=\{F^1,F^2,\ldots,F^m\}$, called *contexts*, a selection probability vector $S=(s^1,s^2,\ldots,s^m)$ over $\mathbf{F}$, a network switching probability q and a random gene perturbation probability p. Random switching and gene perturbation are assumed to be mutually exclusive.

At any time point, with probability q, the network evolution may switch from a constituent Boolean network to another according to the selection probability vector. It is assumed that the probability of switching to a new constituent network is independent of the current network, and in particular the current network may switch to itself. In addition, with probability p, the current state of each gene in the network can be randomly flipped. The PBN is said to be *context-sensitive* if $q<1$, the interpretation being that there are latent variables outside the network whose changes cause the model network to behave stochastically (Brun, Dougherty, & Shmulevich, 2005). If $q=1$, the PBN is said to be *instantaneously-random*, the interpretation being that the uncertainty in the PBN arises from uncertainty in model inference (Shmulevich, Dougherty, Kim, et al., 2002). With sufficiently small perturbation probability p, the long-run behavior of a PBN will reflect the attractor structures within the context Boolean networks (Brun et al., 2005). Throughout this chapter, we assume instantaneously random PBNs.

Transition rules of any instantaneously random PBN can be modeled by a homogeneous Markov chain, whose states are the GAPs of the underlying regulatory network (Shmulevich, Dougherty, Kim, et al., 2002). Let $\{Z_t \in \mathcal{S},\ t=0,1,\ldots\}$ denote the stochastic process of the state of a PBN. For all $t=0,1,\ldots$, originating from state $x \in \mathcal{S}$, the successor state $y \in \mathcal{S}$ is selected randomly according to the TPM $\mathcal{P}$, whose (x,y) element can be calculated readily from the following equation:

$$\left[\mathcal{P}\right]_{xy} = p_{xy} = P(Z_{t+1} = y \mid Z_t = x) = (1-p)^n \sum_{l=1}^{m} s^l 1_{F^l(x)=y} + 1_{D(x,y)\neq 0}(1-p)^{n-D(x,y)} p^{D(x,y)}, \quad (2)$$

where $D(x,y)$ denotes the Hamming distance of x and y in their binary representations. Due to random gene perturbation, the equivalent Markov chain is ergodic and has a unique invariant probability vector, π, equal to its steady-state distribution.

Controlled PBNs

We assume that PBNs admit external control inputs that may take their value from a set of actions specifying the type of intervention on a set of genes (or their products), which directly influences the dynamics of the GRN of interest. We refer to these genes as the *control genes*, and discuss two types of control inputs that have been used in the literature. One can easily extend the theory to other arbitrary control inputs as long as they influence the network dynamics. We assume there is a finite set of control actions, $\mathcal{A}$, where each element $a \in \mathcal{A}$ specifies a control input that effects the expression level of a control gene, g^c, for some $c\in\{1,2,\ldots,n\}$. Since a decision maker might resort to random mechanisms for taking actions, we assume that the actions are permitted to be random. Let $\{A_k \in \mathcal{A}, k=0,1,\ldots\}$ denote the stochastic process of actions taken.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The first, and simpler, class of control inputs are called *binary* inputs, where one allows for choosing decisions from a set $\mathcal{A}=\{0,1\}$ at every transition (or decision epoch) of the process. The control action a=0 results in no-intervention, while a=1 toggles the current state of the control gene. The TPM for the controlled network under a binary input is denoted by $\mathcal{P}(a)$ and defined as

$$\left[\mathcal{P}(a)\right]_{xy} = p_{xy}(a) = P(Z_{t+1} = y \mid Z_t = x, A_t = a), \tag{3}$$

which is the probability of going to y at time t+1 starting from state x and taking action a at time t. Let $\tilde{x}$ have the same binary representation as x except with bit v^c flipped. For a=1, we have

$$\left[\mathcal{P}(1)\right]_{xy} = p_{xy}(1) = P(Z_{t+1} = y \mid Z_t = x, A_t = 1) = P(Z_{t+1} = y \mid Z_t = \tilde{x}) = p_{\tilde{x}y}, \tag{4}$$

where $p_{\tilde{x}y}$ is defined in (2). Thus, the control action a=1 replaces the row corresponding to the state x by the row corresponding to the state $\tilde{x}$ in the original TPM of the underlying Markov chain. For a=0, we simply have $\mathcal{P}(0) = \mathcal{P}$ from (2).

The second class of control inputs are referred to as *persisting effect* inputs, which find application in the intervention models, where the decision maker is not allowed to make decisions at every transition of the network, but only at certain time epochs. Furthermore, once a decision is made it may last for several time epochs until the next decision epoch. To model this effect, one can assume that the external control forcibly sets the expression of the control gene to a certain level. A simple and common choice for $\mathcal{A}$ then would be {0,–1,1}, where 0 indicates no-intervention, and -1 or 1 indicate that the control gene is externally set to 0 or 1, respectively. The network evolves without alteration when a=0, i.e., the transition rules are found from (2). However, for $a\in\{-1,1\}$ the network evolves according to its underlying regulatory functions, but at the same time the control gene is forcibly kept at a fixed state. Hence, depending on the choice of control, some of the states become *accessible* while the others become *inaccessible*, where only transitions between accessible states are permitted. Hence, in the presence of intervention, we will have a PBN with restricted transition rules.

We define $\mathcal{S}_a$ as the set of accessible states when an action $a \in \mathcal{A}$ is in effect. Consequently, the set of inaccessible states for an action $a \in \mathcal{A}$ is $\mathcal{S} \setminus \mathcal{S}_a$. For example, assuming that g^1 is the control gene, which corresponds to the most significant bit in the binary representation of the states, we get

$$\begin{cases} \mathcal{S}_0 = \mathcal{S}, \\ \mathcal{S}_{-1} = \left\{0, 1, 2, \ldots, 2^{n-1} - 1\right\}, \\ \mathcal{S}_1 = \left\{2^{n-1}, 2^{n-1} + 1, \ldots, 2^n - 1\right\}. \end{cases} \tag{5}$$

During a *control phase*, when the network is completely under the influence of a control input $a\in\{-1,1\}$, only transitions between states in $\mathcal{S}_a$ are allowed. Therefore, all the probability mass of possible transitions from an inaccessible state to an accessible state would be absorbed by its corresponding accessible state. We denote the TPM of the controlled network during this phase by $Q(a)$, given by

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\left[\mathcal{Q}(a)\right]_{xy} = q_{xy}(a) = \begin{cases} p_{xy} + p_{x\tilde{y}}, & \text{if } x, y \in \mathcal{S}_a \\ 0, & \text{if } x, y \notin \mathcal{S}_a \end{cases} \tag{6}$$

where $a \in \{-1,1\}$ and p_{xy} is defined in (2). For $a=0$, we have $\mathcal{Q}(0) = \mathcal{P}$.

Next we address two types of transient phases, *uncontrolled-to-controlled* and *controlled-to-uncontrolled*, and find the TPMs during these phases. These transient effects last for only one time epoch until the network enters or leaves the controlled phase. When uncontrolled, the network can start from any original state in $\mathcal{S}$; however, under the control action a it must arrive at a state in $\mathcal{S}_a$ following the transition rules of the controlled network. Using a similar argument as in the control phase, we can calculate the TPM of the network in the uncontrolled to controlled transient phase, denoted by $\mathcal{R}(a)$, as

$$\left[\mathcal{R}(a)\right]_{xy} = r_{xy}(a) = r_{\tilde{x}y}(a) = \begin{cases} p_{xy} + p_{x\tilde{y}}, & \text{if } x, y \in \mathcal{S}_a \\ 0, & \text{if } y \notin \mathcal{S}_a \end{cases} \tag{7}$$

where $a \in \{-1,1\}$ and p_{xy} is defined in (2). For $a=0$, we have $\mathcal{R}(0) = \mathcal{P}$. During the controlled to uncontrolled transient phase, the control action transitions from being effective to ineffective; therefore the state must originate in $\mathcal{S}_a$, depending on the control action taken earlier, but then may transition to any state in the full state space, $\mathcal{S}$. Hence, the network transition probabilities follow that of the original TPM, $\mathcal{P}$.

Cost-Dependent Control

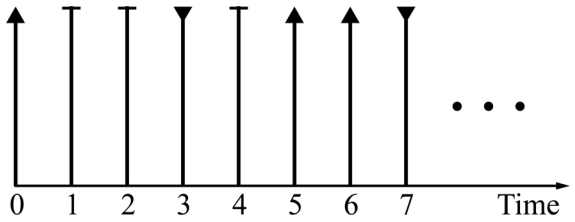
Optimal cost-dependent finite- and infinite-horizon intervention methods assume that there is a bounded cost function associated with every state and control decision pair. Although the choice of cost function is subjective and depending on the treatment goals, typically one assigns non-negative costs to an intervention decision and/or if the next state is an undesirable state.

It has been argued that short term control methods are not always effective as they do not change the steady state behavior of the PBN (Pal et al., 2006); therefore, infinite-horizon methods are preferred, unless one specifically is interested in temporarily changing network dynamics. Optimal infinite-horizon control policies can be formulated relative to two different objective functions: the expected total discounted cost and the expected average cost. Whereas the latter is only concerned with the long-run behavior of the network, the former puts more emphasis on short-term effects of the intervention decisions, although both will change the steady-state behavior of the PBN.

One of the earliest and simplest control methods permits making decisions, at every epoch, to intervene or let the PBN follow its uncontrolled transition rules (Datta et al., 2003; Pal et al., 2006). The optimal control methods for both types of objective functions are stationary, meaning that they are independent of time and only dependent on the state of the network. We refer to this class of control models as *one-transition* models. Figure 1 depicts an example.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Intervention is allowed at each time epoch. Blunt arrows stand for no-intervention; normal arrows stand for up-regulation; reverse arrows stand for down-regulation.

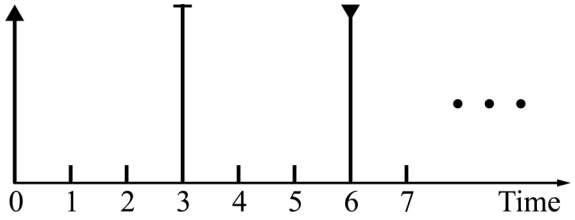


Generalizing the class of one-transition methods in order to account for dosing constraint in chemotherapy, (Vahedi et al., 2009) defines cyclic strategies in which treatment windows are assumed to be every W time epochs. A one-time decision is allowed only at the beginning of each treatment window; thereafter, the network evolves for W–1 epochs without any intervention effects. Similar cost functions as to those in (Pal et al., 2006) are considered and the goal is to minimize the expected total discounted cost. Optimal control methods for this problem are stationary. Figure 2 shows a schematic of this *W-transition* intervention model.

The effect of external intervention, specifically those of antitumor drugs, on tumor cells have been extensively analyzed in the pharmacology literature (Cojocaru & Agur, 1992; de Vladar & González, 2004; Himmelstein & Bischoff, 1973). The pharmacokinetics and pharmacodynamics (PK/PD) of a given drug are separately modeled and then linked together to arrive at a model for the time-course *concentration-effect*. In short, PK uses mathematical models to explain and predict drug concentration in the body (for example, in the blood stream or the tumor site) over time. PD, on the other hand, links the effects of a drug and its mechanism of action to the drug concentration level or the total integrated effective concentration during the course of treatment (Simon & Norton, 2006). It is common to use differential equations (e.g. multi-compartment models) for PK, which in many cases are fairly accurate. However, modeling PD is more challenging as the time-course effects of a drug are not easy to measure. PK/PD linking, with the objective of determining how much variability in drug effectiveness is attributable to measurable drug concentrations, can be done via several functions, including linear models, sigmoid-E_{max} models, etc. (Dayneka, Garg, & Jusko, 1993).

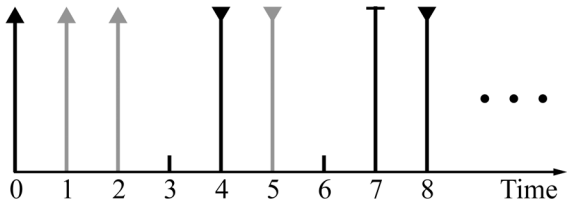
The W-transition policy ignores the drug's duration of action, and the fact that persistent effects exists and must be accounted for in the intervention model. To alleviate this limitation, we build upon recent work for tumor growth inhibition in response to a given drug, where a measure for the duration of ef-

Figure 2. Intervention is only permitted at the beginning of each window of W=3 epochs. Thereafter, the network evolves freely with no intervention for W–1 time epochs. Blunt arrows stand for no-intervention; normal arrows stand for up-regulation; reverse arrows stand for down-regulation.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The drug, once administered, has an action period with a random length, followed by a recovery phase. For example, at epoch 0, the decision is to up-regulate. The drug has a persistent effect of length three epochs (shown by the shaded bars). Similarly, at epoch 4 the decision is to down-regulate, but now the drug effect lasts two epochs. The recovery phase is one epoch. Blunt arrows stand for no-intervention; normal arrows stand for up-regulation; reverse arrows stand for down-regulation.



fective response as a function of parameters is provided (Magni et al., 2008; Magni, Simeoni, Poggesi, Rocchetti, & De Nicolao, 2006). In the new intervention model, the duration of drug action, following each application, is assumed to be random. The randomness is attributed to biological variability among tumor cells, which results in different action periods for a drug. Figure 3 shows an example of this intervention model, referred to as the *random-length* intervention model, where the effect is persistent over a random number of consecutive epochs.

This framework includes several special cases of intervention models:

1. **One-Transition:** Decisions are allowed at every epoch, and drug effect vanishes immediately.
2. ***W*-Transition:** Decisions are allowed at every W epochs, and drug effect vanishes immediately.
3. **Cyclic Control with Fixed-Length Duration of Action:** Decisions are allowed at every W epochs, and drug effect vanishes after $L \leq W$ epochs.
4. **Acyclic Control with Fixed-Length Duration of Action:** Similar to the cyclic model, but when the control decision is *no intervention*, the next decision is made immediately at the next epoch.

We will present the optimal intervention policy for this new intervention model. Our formulation uses the theory of MDPs with partially controlled decision intervals, where given the current state and the action taken, the time interval between successive decisions can be random with a known probability distribution (White, 1978). Similar to the cost function defined for the one-transition model, a predefined cost at each stage is applied when a control action is taken and/or the next state is in the set of undesirable states. The objective is to devise a control strategy that minimizes the expected total discounted cost.

Standard Discounting Cost MDP

We will first review basic MDPs corresponding to the one-transition model. For t=0,1,… denoting the time, let $Z_t \in \mathcal{S}$ and $A_t \in \mathcal{A}$ be a controlled Markov chain, where $\mathcal{S}$ is a finite set of all possible states the Markov chain can occupy, and $\mathcal{A}$ is the set of all possible actions available to the decision maker. The TPM of this chain is defined by

$$p_{xy}(a) = P(Z_{t+1} = y \mid Z_t = x, A_t = a). \tag{8}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A *control policy*, $\mu=\{\mu_t\}$ for $t=0,1,\ldots$ is sequence of instructions for taking actions that steers the dynamics of the process. Let $\{z_t\}$ and $\{a_t\}$ be the sequence of observed states and actions, respectively. We allow for μ_t to be a function of the entire history of the system up to time t, denoted by $h_t = (z_0,a_0, z_1,a_1,\ldots,z_t,a_t)$. Once h_{t-1} is available then some action $a \in \mathcal{A}$ is chosen with probability $\mu_t(a \mid h_{t-1}, z_t)$, where $\mu_t(a \mid h_{t-1}, z_t) \in [0,1]$, and

$$\sum_a \mu_t(a \mid h_{t-1}, z_t) = 1. \tag{9}$$

Let $\mathcal{M}$ be set of all control policies μ. The transition probabilities, $p_{xy}(a)$, the initial state $Z_0=x$ of the Markov chain and any given policy μ determine a unique probability distribution, P_x^μ, over the space of all trajectories of states and actions [derman1970finite]. We denote by E_x^μ the corresponding expectation relative to P_x^μ. Let $g : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \Re$ be the immediate cost function, which is assumed non-negative and bounded for all x, a, and y. Given a discount factor $\lambda\in(0,1)$ and an initial state x, the goal is to find a control policy minimizing the expected total discounted cost defined by

$$J(x,\mu) = \lim_{N\to\infty} E_x^\mu \left\{ \sum_{t=0}^{N} \lambda^t g(Z_t, A_t, Z_{t+1}) \right\}. \tag{10}$$

The discount factor guarantees convergence of (10) (Derman, 1970). The optimal cost function, J^*, is defined as

$$J^*(x) = \min_{\mu\in\mathcal{M}} J(x,\mu), \tag{11}$$

for every $x \in \mathcal{S}$ and a policy μ^* yielding the minimum is called an *optimal policy*. It turns out that one need not consider the entire set $\mathcal{M}$ when searching for the optimal policy; instead it suffices to restrict the search to the set of stationary and deterministic policies, denoted by $\mathcal{M}_D$. Policies, μ_t, in $\mathcal{M}_D$ are the same for all $t=0,1,\ldots$, depend only on the current state, and at each state are equal to one for a single element in $\mathcal{A}$ (Derman, 1970). These three conditions imply that μ^* is a single-valued mapping from $\mathcal{S}$ to $\mathcal{A}$. On the basis of the principles mentioned above, one can formulate requirements to minimize the expected cost over an infinite horizon in terms of a functional equation for the finite horizon case. The resulting optimal policy tells us which action to take at each decision epoch, given the state. The functional equation takes the form,

$$J_{k+1}(x) = \min_a \left\{ \tilde{g}(x,a) + \lambda \sum\nolimits_y p_{xy}(a) J_k(y) \right\}, \tag{12}$$

for $k=0,1,\ldots$ and $x \in \mathcal{S}$, where $\tilde{g}(x,a) = \sum_y p_{xy}(a) g(x,a,y)$. Note that, in the above equation, k does not index forward time; rather, it indicates the number of remaining time epochs. In this new reformulation, $J_{k+1}(x)$ is the minimal expected discounted cost over $k+1$ remaining epochs when the initial state is x and the cost at the final epoch is $J_0(x)$, which is assumed to be bounded for all $x \in \mathcal{S}$.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Starting from state x and by choosing action a, we will move to state y at the beginning of the next epoch with probability $p_{xy}(a)$. Now that the chain is at state y, all decisions for the remaining k epochs are assumed to be optimal. Thus, by definition, the optimal expected discounted cost for these k epochs will be $J_k(y)$ relative to the beginning of the k^{th} epoch. Since this is looking one step ahead, it will be discounted and, therefore, we get a cost of $\lambda J_k(y)$. But at epoch $k+1$ we do not yet know y, so we must take the expected value of $\lambda J_k(y)$ relative to y. The immediate expected cost at the $k+1^{\text{st}}$ epoch is $\tilde{g}(x,a)$ and hence, given x and a, the term on the R.H.S of (12) is the optimal expected discounted cost at epoch $k+1$.

Moving to the infinite-horizon domain, and inspired by (12), one can apply the dynamic programming mapping, $\mathfrak{T}$, transforming any function $J : \mathcal{S} \to \Re$ into a new function $\mathfrak{T}J : \mathcal{S} \to \Re$, defined by

$$(\mathfrak{T}J)(x) = \min_a \left\{ \tilde{g}(x,a) + \lambda \sum\nolimits_y p_{xy}(a) J(y) \right\}, \tag{13}$$

for $x \in \mathcal{S}$. We can determine the optimal policy with the help of the following theorems, resulting in an iterative algorithm called *value iteration*. Proofs of these statements can be found in (Bertsekas, 2005).

Theorem 1.1 (Convergence): For any bounded function $J : \mathcal{S} \to \Re$ and all $x \in \mathcal{S}$, the optimal cost function satisfies

$$J^*(x) = \lim_{N \to \infty} (\mathfrak{T}^N J)(x), \tag{14}$$

where $\mathfrak{T}^N$ is the composition of the mapping $\mathfrak{T}$ with itself N times.

Theorem 1.2 (Bellman's Optimality Equation): For all $x \in \mathcal{S}$, the optimal cost function J^* satisfies

$$J^*(x) = \min_a \left\{ \tilde{g}(x,a) + \lambda \sum\nolimits_y p_{xy}(a) J^*(y) \right\}. \tag{15}$$

Furthermore, J^* is the unique solution of this equation in the set of all bounded functions.

Theorem 1.3 (Necessary and Sufficient Condition for Optimality): A stationary and deterministic policy μ is optimal if and only if $\mu(x)$ attains the minimum in Bellman's optimality equation for every $x \in \mathcal{S}$.

The first part of Theorem 1.2 says that $J^* = \mathbf{T}J^*$. In other words, J^* is the fixed point of $\mathbf{T}$. Theorem 1.1 states that J^* can be iteratively found by successively computing $\mathfrak{T}J, \mathfrak{T}^2 J, \ldots$ for any bounded initial cost function J. The convergence of this algorithm is guaranteed due to discounting, where the limit is J^*. Hence, the value iteration algorithm can be continued until some stopping criterion is reached. The last theorem assures that the resulting optimal policy is stationary and deterministic.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Discounting Cost MDP with Partially Controlled Decision Intervals

In a standard MDP, the time index and decision epochs are the same. However, for a MDP with partially controlled decision intervals decision epochs might include multiple time units (White, 1978). Define $\mathcal{L} = \{1, 2, \ldots, K\}$, for a fixed $K \geq 1$, to be the largest time interval between two consecutive decisions epochs. Let t=0,1,… index the decision epochs. We denote by $L_t \in \mathcal{L}$ and $\{l_t\}$ the sequences of random variables representing the length of the time intervals between consecutive decision epochs and the corresponding observed values, respectively. We assume $L_0 \equiv 0$. The transition probabilities for the Markov chain with partially controlled decision intervals are characterized by

$$p_{xy}^{l}(a) = P(L_{t+1} = l, Z_{t+1} = y \mid Z_t = x, A_t = a), \tag{16}$$

for any t=0,1,…, where $Z_t \in \mathcal{S}$ is the sequence of random states, $A_t \in \mathcal{A}$ the sequence of actions, and $\{z_t\}$ and $\{a_t\}$ are the corresponding sequences of observed states and actions, respectively. $p_{xy}^{l}(a)$ is the probability that if we start at a state $x \in \mathcal{S}$ and the decision maker takes an action $a \in \mathcal{A}$, the next call for a decision will take place $l \in \mathcal{L}$ time epochs later and we arrive at state $y \in \mathcal{S}$.

We define a *control policy* μ to be a sequence $\{\mu_t\}$ of instructions for taking actions at each decision epoch t. Each μ_t may be a function of the entire history of the system up to decision epoch t, which we define by

$$h_t = (l, z_0, a_0, l_1, z_1, a_1, \ldots, l_t, z_t, a_t). \tag{17}$$

Once we have h_{t-1}, we will choose an action $a \in \mathcal{A}$ with probability $\mu_t(a \mid h_{t-1}, l_t, z_t) \in [0,1]$, such that

$$\sum_a \mu_t(a \mid h_{t-1}, l_t, z_t) = 1. \tag{18}$$

Denote the set of all control policies μ by $\mathcal{M}$. Similar to the standard MDP, the transition probabilities, $p_{xy}^{l}(a)$, the initial state, $Z_0 = x$, and the policy, μ, determine P_x^{μ} over the space of all trajectories of the states, actions and decision epochs. Let E_x^{μ} denote the corresponding expectation operator. We define $g(x, a, l, y) : \mathcal{S} \times \mathcal{A} \times \mathcal{L} \times \mathcal{S} \to \Re$ to be the immediate cost function over a decision epoch of length l when a transition from state x under control a to state y is made in l subsequent periods. We assume g is non-negative and bounded for all tuples (x,a,l,y). The objective function is the expected total discounted cost, defined by

$$J(x, \mu) = \lim_{N \to \infty} E_x^{\mu} \left\{ \sum_{t=0}^{N} \lambda^{\left(\sum_{m=0}^{t} L_m\right)} g(Z_t, A_t, L_{t+1}, Z_{t+1}) \right\}, \tag{19}$$

for every state $x \in \mathcal{S}$ and $\mu \in \mathcal{M}$, and the goal is to minimize it over the space of policies $\mathcal{M}$. (White, 1978) formulated the following equivalent functional equation for $k \geq K-1$ and all $x \in \mathcal{S}$:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$J_{k+1}(x) = \min_a \left\{ \tilde{g}(x,a) + \sum_y \sum_l \lambda^l p^l_{xy}(a) J_{k+1-l}(y) \right\}, \tag{20}$$

$$\tilde{g}(x,a) = \sum_y \sum_l p^l_{xy}(a) g(x,a,l,y), \tag{21}$$

and $\lambda \in (0,1)$. The functions $J_k(x)$ are assumed given and bounded for all $1 \le k \le K-2$ and $x \in \mathcal{S}$. In, $J_{k+1}(x)$ is the minimal expected discounted cost beginning in state x with $k+1$ remaining epochs and $\tilde{g}(x,a)$ is the immediate expected cost at state x and for control action a. If the next decision epoch happens l time epochs later and leads to arrival at state y the minimal expected discounted cost (relative to that point in time) is $J_{k+1-l}(y)$, which when discounted to the present time by a factor λ^l, and when averaged over the joint distribution of y and l, gives the required result by minimizing over all actions $a \in \mathcal{A}$ (White, 1978).

We define the dynamic programming mapping, $\mathfrak{T}$, by

$$(\mathfrak{T}J)(x) = \min_a \left\{ \tilde{g}(x,a) + \sum_y \sum_l \lambda^l p^l_{xy}(a) J(y) \right\}, \tag{22}$$

for all $x \in \mathcal{S}$. Finding the fixed point of the mapping $\mathfrak{T}$, and the minimizing control action at each state x, gives the control policy minimizing the objective function (19). The following three theorems provide the solution to this problem.

Theorem 1.4 (Convergence): For any bounded function $J : \mathcal{S} \to \mathbb{R}$ and all $x \in \mathcal{S}$, the optimal cost function satisfies

$$J^*(x) = \lim_{N \to \infty} (\mathfrak{T}^N J)(x). \tag{23}$$

Theorem 1.5 (Bellman's Optimality Equation): For all $x \in \mathcal{S}$, the optimal cost function J^* satisfies

$$J^*(x) = \min_a \left\{ \tilde{g}(x,a) + \sum_y \sum_l \lambda^l p^l_{xy}(a) J^*(y) \right\}. \tag{24}$$

Equivalently, $J^* = \mathfrak{T}J^*$, and J^* is the unique solution of this equation in the set of all bounded functions.

Theorem 1.6 (Necessary and Sufficient Condition for Optimality): A stationary and deterministic policy μ is optimal if and only if $\mu(x)$ attains the minimum in Bellman's optimality equation for every $x \in \mathcal{S}$.

Proofs of these statements can be found in (Yousefi, Datta, & Dougherty, 2012). Similar to the standard MDP, the value iteration algorithm can be used to find the optimal policy.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Random-Length Intervention Model

Successful treatment, both at the individual level or the population level, requires modeling uncertainties in the underlying dynamical system corresponding to the tumor, its interaction with the drug, and the costs associated with the treatment (Heppner, 1984). For each treatment cycle, samples must be taken from patients to determine the state of the tumor. Depending on the state, a decision must be made as to what type of intervention is necessary. An intervention at the beginning of a treatment cycle usually implies administering an antitumor drug that initiates the process of tumor growth inhibition.

As a special case of the finite-dimensional growth model, (Magni et al., 2008, 2006) propose a semi-empirical tumor growth inhibition model that exhibits good prediction for tumor growth for *in vivo* experiments. Once the effect of the drug on the dynamics of an unperturbed tumor is modeled, under some simplifying assumptions, they show that, in a single treatment window, tumor growth is delayed. The extent of this effect approximately depends on the proportionality constant of a *damage* function, which models the desirable effect of the drug on the tumor cells, the tumor growth rate at early stages of development, and the integral of the drug concentration-time curve. We denote by T_e the length of the drug effect on tumor cells over a single treatment window. Taking into account the variability of response to the antitumor drug within a population, we assume a four-parameter beta distribution for T_e,

$$T_e \sim \mathrm{Beta}(\alpha, \beta, \underline{T}_e, \overline{T}_e), \tag{25}$$

defined on the interval $(\underline{T}_e, \overline{T}_e)$, where α,β>0 are two real numbers characterizing the shape of this distribution. The parameters of this distribution may be set after consulting with pharmacologists and drug developers.

Assuming PBNs as our reference model for the overall dynamical behavior of tumor cells in the population and external control inputs with persisting effect with $\mathcal{A}=\{0,1,-1\}$, we denote by $T_e(a)$ the length of the window during which the corresponding control gene, directly affected by the drug, will be invariant depending on the type of intervention action $a \in \mathcal{A}$. We also assume that there is a fixed-length rest window of length $T_r(a)$, following $T_e(a)$, to let the side effects of the drug gradually subside. In addition, in the case of no-intervention, i.e. a=0, the next decision will be called after $T_r(0)$ units of time have elapsed since $T_e(0)$ is zero. Note that transitions in a PBN occur in discrete time, while $T_e(a)$ and $T_r(a)$ are continuous-time quantities. A simple resolution would be to construct a probability mass function over a discrete set of levels for these variables. We denote the discretized variables by $\overline{T_e(a)}$ and $\overline{T_r(a)}$.

We define a cost function over the set of control actions and the set of states a PBN may occupy. The total cost for the entire course of treatment is obtained by accumulating the discounted cost incurred at each transition of the network. The discounted cost emphasizes that obtaining a response at earlier stages is favored over later stages. The cost functions directly reflect our objectives; hence, they should be assigned by physicians in conjunction with pharmacologists according to their understanding of the treatment objectives.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

To solve the optimal intervention problem for this particular model, we need only calculate $p_{xy}^{l}(a)$ and $\tilde{g}(x,a)$ for every *x,y,a,* and *l*, and use the mathematical machinery for the MDPs with partially controlled decision intervals. To this end, we first define $\mathcal{L}_{w}(a)=\left\{0,1,2,\ldots,W(a)\right\}$, where *W*(*a*) denotes the maximum length of response to the intervention action *a*, and we define

$$\mathcal{L}(a)=\left\{l: l=\overline{T_{e}(a)}+\overline{T_{r}(a)}\,\text{and}\,\overline{T_{e}(a)}\in\mathcal{L}_{w}(a)\right\}. \tag{26}$$

The probability mass function of $\overline{T_{e}(a)}$ is defined as $P_{l_{w}}(a)=P\left(\overline{T_{e}(a)}=l_{w}\right)$ for all $l_{w}\in\mathcal{L}_{w}(a)$ and $a\in\mathcal{A}$. Note $P_{l_{w}}(a)\geq 0$ and $\sum_{l_{w}}P_{l_{w}}(a)=1$. Thus, the total length of each treatment cycle (drug action and rest together) is a discrete random variable whose probability mass function is $P_{l}(a)=P_{l_{w}}(a)$, for every $l\in\mathcal{L}(a)$ and $a\in\mathcal{A}$.

One structural constraint of the problem is that if no intervention occurs, i.e. $\overline{T_{e}(0)}=0$, then the next decision epoch will be called deterministically $\overline{T_{r}(0)}$ time epochs later on. Thus, we have $P_{l}(0)$=1 for $l=\overline{T_{r}(0)}$. Another structural constraint is that if any intervention occurs we have $\overline{T_{e}(a)}\geq 1$, meaning that $P_{l}(a)$=0 for $l=\overline{T_{r}(a)}$ when $a\in\mathcal{A}\setminus\{0\}$. Therefore,

$$p_{\mathbf{ij}}^{l}(0)=\begin{cases}p_{\mathbf{ij}}^{\left(\overline{T_{r}(0)}\right)}, & \text{if } l=\overline{T_{r}(0)}\\ 0, & \text{otherwise}\end{cases} \tag{27}$$

$$p_{xy}^{l}(a)=P_{l}(a)\sum_{x'}r_{xx'}(a)\sum_{x''}q_{x'x''}^{\left(l-\overline{T_{r}(a)}-1\right)}(a)p_{x''y}^{\left(\overline{T_{r}(a)}\right)}, \tag{28}$$

for all $a\in\mathcal{A}\setminus\{0\}$ and all $l\in\mathcal{L}(a)$. The terms p_{xy}, q_{xy}, and r_{xy} are found from (2), (6) and (7), respectively, and a superscript (*n*), as in $p_{xy}^{(n)}$, denotes the transition probability of going from state *x* to *y* in *n* epochs.

In our setting, the immediate cost function $\tilde{g}(x,a,l,y)$ is also an expected discounted cost accumulated over the entire window. Let $g_{xy}(a)$ denote the one-transition cost function, which depends on the current state *x*, next state *y* and control *a*. We assume that $g_{xy}(a)$ is non-negative and bounded for all $x,y\in\mathcal{S}$ and $a\in\mathcal{A}$, and postulate the following structure for it:

$$g_{xy}(a)=\begin{cases}C+c(a), & \text{if } y\in\mathcal{U}\\ c(a), & \text{if } y\in\mathcal{D}\end{cases} \tag{29}$$

where *c*(*a*)=*c* for all $a\in\mathcal{A}\setminus\{0\}$ and zero otherwise. It is straightforward to show that

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\tilde{g}(x,a)=\begin{cases}\sum\limits_{m=0}^{\overline{T_r(0)}-1}\lambda^m\sum\limits_y p_{xy}^{(m+1)}g_{xy}(a) & \text{for } a=0\\ \sum\limits_{l\in\mathcal{L}(a)}P_l(a)\tilde{g}(x,a,l) & \text{for } a\in\mathcal{A}\setminus\{0\}\end{cases} \tag{30}$$

$$\begin{aligned}\tilde{g}(x,a,l)=&\sum_y r_{xy}(a)g_{xy}(a)+\sum_{m=1}^{l-\overline{T_r(a)}-1}\lambda^m\sum_{x'}r_{xx'}(a)\sum_y q_{x'y}^{(m)}(a)g_{x'y}(a)\\&+\sum_{m=l-\overline{T_r(a)}}^{l-1}\lambda^m\sum_{x'}r_{xx'}(a)\sum_{x''}q_{x'x''}^{(l-\overline{T_r(a)}-1)}(a)\sum_y p_{x''y}^{(m-l+\overline{T_r(a)}+1)}g_{x''y}(0),\end{aligned} \tag{31}$$

for all $l\in\mathcal{L}(a)$. Substituting $p_{xy}^{l}(a)$ and $\tilde{g}(x,a)$ from, and, respectively, into and using the value iteration algorithm one can find the optimal control policy.

Simulation on Networks

To illustrate the effectiveness of our approach, we run several simulations on both synthetically generated PBNs and a PBN corresponding to a mammalian cell cycle network. Throughout, we choose the following values for the parameters:

$$\text{Cost function parameters:}\begin{cases}\lambda=0.95\\ C=2\\ c=0.1\end{cases}\quad\cdots\quad\text{Drug effect parameters:}\begin{cases}W(-1)=W(1)=W\\ W=10\\ \overline{T_r(a)}\equiv\overline{T_r}\in\{1,2,5,10,20\}\\ T_e(-1)=T_e(1)=T_e\\ T_e\sim\text{Beta}(10,30,0,10)\end{cases}$$

We also discretize the beta distributions into ten levels. The resulting discrete distributions are shown in Figure 4.

We evaluate six intervention strategies, each based on different assumptions and constraints:

M1 (One-Transition): Decisions are made at each time epoch and the drug duration of effect is assumed to be 1;

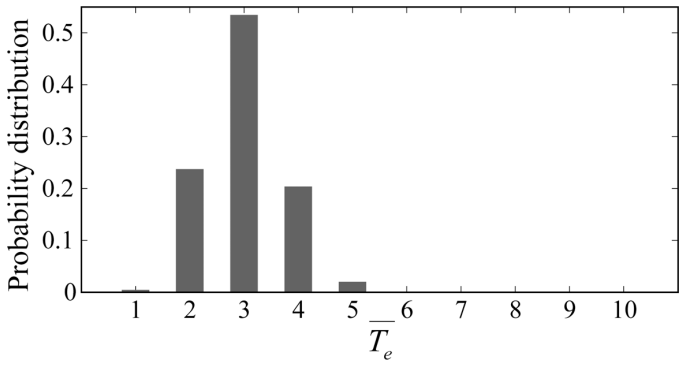
M2 (Cyclic $(W+\overline{T_r})$-Transition without Drug Effect): Decisions are made at every $(W+\overline{T_r})$ time epochs and the drug duration of effect is assumed to be 1;

M3 (Acyclic $(W+\overline{T_r})$-Transition without Drug Effect): Decisions are made at every $(W+\overline{T_r})$ time epochs if $a\in\{-1,1\}$ and $\overline{T_r}$ time epochs otherwise. The drug duration of effect is assumed to be 1;

M4 (Cyclic $(W+\overline{T_r})$-Transition with Drug Effect): Similar to M2 but the drug response is modeled with fixed length of *W*;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Quantized beta distributions for $\overline{T}_e$ with T_e~Beta(10,30,0,10)



M5 (Acyclic $(W + \overline{T}_r)$ -Transition with Drug Effect): Similar to M3 but the drug response is modeled with fixed length of *W*;

M6 (Random-Length Drug Effect): The drug is correctly modeled with a random-length effect, and we allow for a rest period of length $\overline{T}_r$.

We refer to the optimal policies derived for each model by the name of that model. Once control policies are found, we apply them with the random-length drug effect (M6) serving as the true model. Note that M1-M5 policies are suboptimal for this model. Hence, our simulation will show how much, relative to the suboptimal policies, one can improve upon when using the optimal policy.

Synthetic Networks

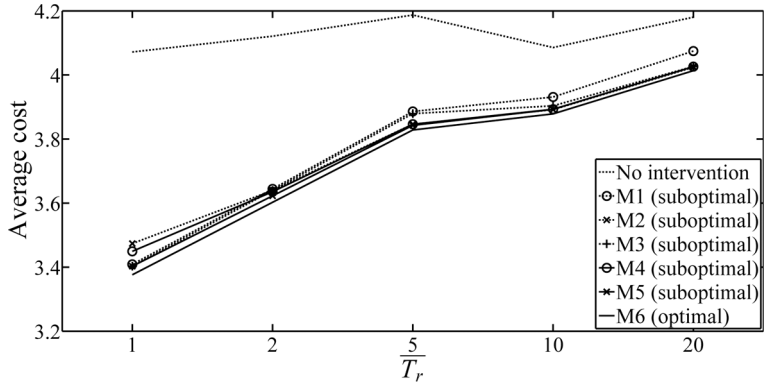
We consider instantaneously random PBNs with seven genes. We also assume each PBN consists of four equally likely constituent Boolean networks with the maximum number of predictors for each Boolean function set to three. When generating Boolean updating functions for each gene, we set the output of each Boolean function, for each input, to one with a small probability, which itself is drawn randomly from a beta distribution with mean 0.1 and standard deviation 0.01. The gene perturbation probability is set to p=0.01. Once BNs are generated, we compute the TPM of each PBN. We define undesirable states, $\mathcal{U}$, as those for which the least significant bit in their binary representation is one, and assume that the control gene is the most significant bit of states.

For each set of simulation parameters, we generate 2900 random PBNs. For each randomly generated PBN in a set of simulations, we find the optimal policies for M1-M6, implement them on the M6 model by simulating the intervention processes over a time-course of length 1500 time epochs, and accumulate the discounted cost incurred at each state-control pair observed. We repeat this process 50000 times, each with a random (uniformly generated) initial state, and compute the expected total discounted cost.

Figure 5 shows the effect of different values for $\overline{T}_r$ on the expected total discounted cost, averaged over 2900 random networks. In this case, $\overline{T}_e$ will most likely take on values from {2,3,4} as shown in Figure 4 which indicates a relatively small variability in the length of drug response. We observe that M6 is optimal on average with a small margin compared to the suboptimal policies.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Average cost of different policies across 2900 random PBNs for $\overline{T}_r \in \{1, 2, 5, 10, 20\}$



Although suboptimal policies appear to perform similarly to the optimal one on average, there are networks for which the optimal policy clearly outperforms the suboptimal ones. We quantify the amount of reduction in the cost, for each random network, achieved by the optimal intervention policy relative to the suboptimal ones via the following formula:

$$\Delta(\mu) = 100 \times \frac{J(\mu) - J(\mu^{\mathrm{M6}})}{J(\mu)}, \tag{32}$$

where $J(\mu)$ denotes the expected total discounted cost when a policy μ is used and $J(\mu^{\mathrm{M6}})$ denotes the expected total discounted cost when under the optimal policy, M6. Because the networks are randomly generated Δ will also be random. We graph the empirical complementary cumulative distribution function (CDF) of Δ for different suboptimal polices when $\overline{T}_r \in \{1, 5, 20\}$.

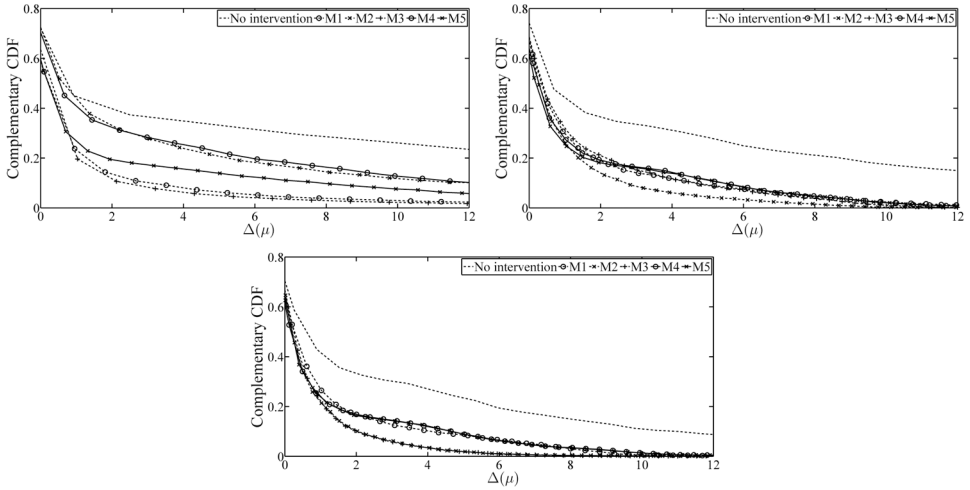
The complementary CDF graph in Figure 6 reveals a high probability that the optimal policy noticeably outperforms others. Interestingly, we observe very little improvement over the M2 and M3 policies for larger $\overline{T}_r$. In this case, $\overline{T}_e$ is likely to be much smaller than $\overline{T}_r$, so that the policies resulting from M2 and M3 may induce similar cost compared to M6.

Real Network

This network models normal biological processes that occur during the mammalian cell cycle, whose disruption can lead to uncontrolled cell proliferation. It was originally constructed in (Faure, Naldi, Chaouiya, & Thieffry, 2006) to provide a logical framework for comparing different strategies to explore the dynamical properties of the core mechanisms controlling the mammalian cell cycle. The network has ten genes (or proteins), CycD, Rb, p27, E2F, CycE, CycA, Cdc20, Cdh1, UbcH10 and CycB, and the regulatory relationships are assumed to be Boolean. Table 1 summarizes these regulatory functions. Three key components in this network are Cyclin D (CycD), retinoblastoma (Rb) and p27. Extracellular signals, under normal conditions, coordinate cell division with overall growth by controlling the activation of CycD. The tumor suppressor protein, Rb, plays a central role in the negative control of the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Empirical complementary CDF of the improvement over different intervention policies across 2900 random PBNs: (Top-left) $\overline{T}_r = 1$ *; (Top-right)* $\overline{T}_r = 5$ *; (Bottom)* $\overline{T}_r = 20$



cell cycle and in tumor progression (Giacinti & Giordano, 2006). Rb is active in the absence of certain cyclin-dependent kinases. When present, however, these cyclins inhibit Rb through phosphorylation. p27 is also active in the absence of the cyclins. An active p27 blocks the action of CycE or CycA and, hence, Rb can also be active even if the cyclins are present, resulting in a mechanism that stops uncontrolled cell division.

This Boolean network can also be used to study the consequences of various mutations. For instance, a possible mutation of p27 (always OFF) may introduce detrimental effects on a cell's normal regulation (Faure et al., 2006). A situation can arise, where both CycD and Rb might be inactive and cell division occurs repeatedly in the absence of any growth factor. This suggests labeling the states in which both Rb and CycD are down-regulated, under a mutated p27, as undesirable. Table 2 presents the Boolean functions of this mutated cell cycle network. The network consists of nine genes, CycD, Rb, E2F, CycE, CycA, Cdc20, Cdh1, UbcH10 and CycB, ordered from the most significant bit to the least significant bit in the binary representation. Depending on the state of the extracellular signal, which determines CycD being ON or OFF, we obtain two contexts. We first build an instantaneously random PBN where these two contexts described above are equally likely. We assume that the gene perturbation probability is 0.01, and define $\mathcal{U}$ to be those states for which CycD and Rb are simultaneously down-regulated. This reflects the undesirability of cell growth in the absence of growth factors. We also choose E2F as the control gene.

Similar to the synthetic networks, we find the M1-M6 policies and apply them to the M6 intervention model. We implement a similar procedure for estimating the expected total discounted cost, but with a small difference where we iterate the time course simulation 250000 times, each with a random initial state, and set the length of the time course simulation to 2500 time epochs. Figure 7 shows performance comparison results for each intervention policy based on the expected total discounted cost for different sets of parameters. The performance results are consistent with the results obtained from simulations on the synthetic networks.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Boolean network of a wild-type mammalian cell cycle

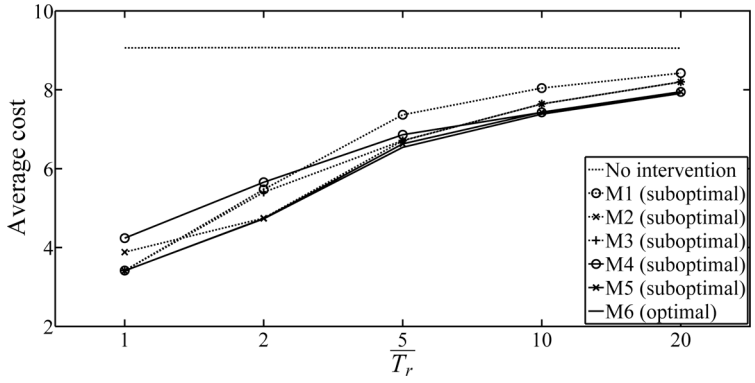
Node Name	Gene Name	Predictor Function
v^1	CycD	Extracellular signal
v^2	Rb	$\left(\overline{\text{CycD}}\ \textbf{AND}\ \overline{\text{CycE}}\ \textbf{AND}\ \overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)\ \textbf{OR}\ \left(\text{p27}\ \textbf{AND}\ \overline{\text{CycD}}\ \textbf{AND}\ \overline{\text{CycB}}\right)$
v^3	E2F	$\left(\overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)\ \textbf{OR}\ \left(\text{p27}\ \textbf{AND}\ \overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{CycB}}\right)$
v^4	CycE	$\text{E2F}\ \textbf{AND}\ \overline{\text{Rb}}$
v^5	CycA	$\left(\text{E2F}\ \textbf{AND}\ \overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{Cdc20}}\ \textbf{AND}\ \left(\overline{\text{Cdh1}\ \textbf{AND}\ \text{UbcH10}}\right)\right)\ \textbf{OR}\ \left(\text{CycA}\ \textbf{AND}\ \overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{Cdc20}}\ \textbf{AND}\ \left(\overline{\text{Cdh1}\ \textbf{AND}\ \text{UbcH10}}\right)\right)$
v^6	p27	$\left(\overline{\text{CycD}}\ \textbf{AND}\ \overline{\text{CycE}}\ \textbf{AND}\ \overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)\ \textbf{OR}\ \left(\text{p27}\ \textbf{AND}\ \overline{\text{CycE}\ \textbf{AND}\ \text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\ \textbf{AND}\ \overline{\text{CycD}}\right)$
v^7	Cdc20	CycB
v^8	Cdh1	$\left(\overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)\ \textbf{OR}\ \text{Cdc20}$
v^9	UbcH10	$\left(\text{Cdh1}\ \textbf{AND}\ \text{UbcH10}\ \textbf{AND}\ \left(\text{Cdc20}\ \textbf{OR}\ \text{CycA}\ \textbf{OR}\ \text{CycB}\right)\right)\textbf{OR}\ \overline{\text{Cdh1}}$
v^{10}	CycB	$\overline{\text{Cdc20}}\ \textbf{AND}\ \overline{\text{Cdh1}}$

Table 2. Boolean network of a mutated mammalian cell cycle network

Node Name	Gene Name	Predictor Function
v^1	CycD	Extracellular signal
v^2	Rb	$\left(\overline{\text{CycD}}\ \textbf{AND}\ \overline{\text{CycE}}\ \textbf{AND}\ \overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)$
v^3	E2F	$\left(\overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)$
v^4	CycE	$\text{E2F}\ \textbf{AND}\ \overline{\text{Rb}}$
v^5	CycA	$\left(\text{E2F}\ \textbf{AND}\ \overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{Cdc20}}\ \textbf{AND}\ \left(\overline{\text{Cdh1}\ \textbf{AND}\ \text{UbcH10}}\right)\right)\ \textbf{OR}\ \left(\text{CycA}\ \textbf{AND}\ \overline{\text{Rb}}\ \textbf{AND}\ \overline{\text{Cdc20}}\ \textbf{AND}\ \left(\overline{\text{Cdh1}\ \textbf{AND}\ \text{UbcH10}}\right)\right)$
v^6	Cdc20	CycB
v^7	Cdh1	$\left(\overline{\text{CycA}}\ \textbf{AND}\ \overline{\text{CycB}}\right)\ \textbf{OR}\ \text{Cdc20}$
v^8	UbcH10	$\left(\text{Cdh1}\ \textbf{AND}\ \text{UbcH10}\ \textbf{AND}\ \left(\text{Cdc20}\ \textbf{OR}\ \text{CycA}\ \textbf{OR}\ \text{CycB}\right)\right)\textbf{OR}\ \overline{\text{Cdh1}}$
v^9	CycB	$\overline{\text{Cdc20}}\ \textbf{AND}\ \overline{\text{Cdh1}}$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Average total discounted cost of control policies derived from different intervention models for the cell cycle network and $\overline{T}_r \in \{1, 2, 5, 10, 20\}$



Cost-Independent Control

In the previous section, we assigned a large cost to undesirable states and aimed to minimize the long-run discounted cost of being in these states. In systems biology, it is common to model phenotypes as the steady-state behavior of the underlying regulatory network. Thus, from a practical perspective, in many cases our primary interest may be to optimize the long-run occupation of undesirable states, and not to minimize a subjective cost function. Direct optimization has been addressed to some extent in (Qian & Dougherty, 2012; Qian, Ivanov, Ghaffari, & Dougherty, 2009), where the authors propose several classes of greedy stationary policies to reduce the mass corresponding to undesirable states and increase the mass corresponding to desirable states in the steady state distribution, subject to some constraints.

In this section, we will review these suboptimal methods and present an optimal solution to the maximal phenotype alteration problem. Our solution provides an optimal policy with a computational complexity equivalent to solving a linear program (LP) optimization problem. We also address a generalization of this optimization procedure, where the amount of mass moved to some ambiguous states may be constrained. This constrained optimization is especially important in therapeutic methods since one would like to avoid introducing new probability mass to states associated with unknown phenotypes as much as possible (Qian & Dougherty, 2012).

Suboptimal Phenotype Alteration

Consider an instantaneously random PBN with a single control gene g^c and a binary control input. One of the earliest suboptimal methods for maximally altering the long-run behavior of a PBN is a greedy algorithm based on *mean-first-passage-time (MFPT)* (Vahedi, Faryabi, Chamberland, Datta, & Dougherty, 2008), which locates a policy in $\mathcal{M}_{\mathrm{D}}$ based on the heuristic that reaching desirable states and leaving undesirable states as soon as possible is preferable. One of the advantages of this algorithm is that it is model-free and can be directly designed from time-course data without having to infer the TPM of the network.

Suppose that up-regulation of the gene g^1 is undesirable. The TPM of the uncontrolled PBN can be partitioned into four sub-matrices as follows:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\mathcal{P}(0) = \begin{pmatrix} \mathcal{P}_{\mathcal{D}\mathcal{D}} \mathcal{P}_{\mathcal{D}\mathcal{U}} \\ \mathcal{P}_{\mathcal{U}\mathcal{D}} \mathcal{P}_{\mathcal{U}\mathcal{U}} \end{pmatrix}. \tag{33}$$

One can compute the MFPS vectors $MFPT_{\mathcal{D}}$ (the MFPTs from each state in $\mathcal{U}$ to any state in $\mathcal{D}$) and $MFPT_{\mathcal{U}}$ (the MFPTs from each state in $\mathcal{D}$ to any state in $\mathcal{U}$) by solving the following system of linear equations (Kemény & Snell, 1960):

$$\begin{cases} MFPT_{\mathcal{D}} = e + \mathcal{P}_{\mathcal{D}\mathcal{D}} MFPT_{\mathcal{D}} \\ MFPT_{\mathcal{U}} = e + \mathcal{P}_{\mathcal{U}\mathcal{U}} MFPT_{\mathcal{U}} \end{cases} \tag{34}$$

where *e* is a column vector of 1's with the appropriate length (with *n* genes, *e* has 2^{n-1} elements). We denote by $\mu_{\text{MFPT}} : \mathcal{S} \to \mathcal{A}$ the control policy derived from the following procedure: choose a number γ. For every $x \in \mathcal{U}$, if $MFPT_{\mathcal{D}}(x) - MFPT_{\mathcal{D}}(\tilde{x}) > \gamma$, then $\mu_{\text{MFPT}}(x)=1$; otherwise, $\mu_{\text{MFPT}}(x)=0$. Using this rule, one reduces the expected time to reach the desirable states. Analogously, for every $x \in \mathcal{D}$, if $MFPT_{\mathcal{U}}(\tilde{x}) - MFPT_{\mathcal{U}}(x) > \gamma$, then $\mu_{\text{MFPT}}(x)=1$; otherwise, $\mu_{\text{MFPT}}(x)=0$. Thus, the expected time until a passage to undesirable states is increased. The tuning parameter γ is set to a large when the ratio of the cost of control to the cost of the undesirable states is higher so that the control is applied less frequently; otherwise, under no constraints on the control frequency, we set γ=0 (Vahedi et al., 2008).

Another greedy method for solving this problem is called a *basin of attraction (BOA)* control policy, which belongs to the set of stationary and deterministic policies (Qian et al., 2009). For any state *x*, let $\mathcal{L}^j(x)$ be the set of attractors for the basin containing a state *x* in the *j*th constituent Boolean network, keeping in mind that a state *x* belongs to exactly one basin of attraction in each constituent Boolean network. Let

$$\mathcal{B}(x) = \bigcup_{j=1}^{m} \mathcal{L}^j(x) \tag{35}$$

be the union of attractors for the basins of *x* over all of the constituent Boolean networks. For each constituent Boolean network, we calculate the minimal distance of each state in $\mathcal{S}$ to states in $\mathcal{D}$ or in $\mathcal{U}$, and denote their weighted average, with respect to the Boolean networks selection probabilities, by $d_{\mathcal{D}}$ and $d_{\mathcal{U}}$, respectively (Qian et al., 2009).

A BOA control policy, $\mu_{\text{BOA}} : \mathcal{S} \to \mathcal{A}$, is derived via the following algorithm for every $x \in \mathcal{S}$. Suppose $x \in \mathcal{U}$. If $\mathcal{B}(x)$ does not contains any desirable attractors and at the same time $\mathcal{B}(\tilde{x})$ contains desirable attractors, then $\mu_{\text{BOA}}(x)=1$. By this, the network moves to $\tilde{x}$ in which the likelihood of arriving into desirable attractors is increased. If both or neither of $\mathcal{B}(x)$ and $\mathcal{B}(\tilde{x})$ contain desirable attractors, then $\mu_{\text{BOA}}(x)=1$ if $d_{\mathcal{D}}(x) > d_{\mathcal{D}}(\tilde{x})$, $\mu_{\text{BOA}}(x)=0$ if $d_{\mathcal{D}}(x) < d_{\mathcal{D}}(\tilde{x})$, or otherwise, $\mu_{\text{BOA}}(x)=0$. Now, suppose $x \in \mathcal{D}$. If $\mathcal{B}(x)$ contains undesirable attractors and at the same time $\mathcal{B}(\tilde{x})$ does not contain any undesirable attractors, $\mu_{\text{BOA}}(x)=1$, so that the chance of entering undesirable attractors is reduced. If both or neither of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$\mathcal{B}(x)$ and $\mathcal{B}(\tilde{x})$ contain undesirable attractors, then $\mu_{\text{BOA}}(x)=1$ if $d_{\mathcal{U}}(\tilde{x}) > d_{\mathcal{U}}(x)$, or $\mu_{\text{BOA}}(x)=0$ if $d_{\mathcal{U}}(\tilde{x}) < d_{\mathcal{U}}(x)$. Otherwise, $\mu_{\text{BOA}}(x)=0$ (Qian et al., 2009).

A third algorithm, the *steady-state distribution (SSD)* control policy, is based on the theory of rank-one perturbations of regulatory functions, where one can readily calculate the exact shifted stationary mass (Qian & Dougherty, 2008; Qian et al., 2009). Let π and $\mathcal{Z} = [I - \mathcal{P}(0) + e\pi]^{-1}$ be the steady-state row vector and the fundamental matrix of the unperturbed network, respectively, where *I* is the identity matrix *e* is a column vector of 1's with the appropriate dimensions. The perturbed steady-state distribution can be computed by:

$$\hat{\pi}^{x} = \pi - \frac{\pi_{x}(\mathcal{P}_{x}(0) - \mathcal{P}_{\tilde{x}}(0))\mathcal{Z}}{1 + (\mathcal{P}_{x}(0) - \mathcal{P}_{\tilde{x}}(0))\mathcal{Z}^{x}}, \tag{36}$$

where $\mathcal{P}_{x}$ is a row in $\mathcal{P}$ corresponding to the states *x*, $\mathcal{Z}^{x}$ is a column in $\mathcal{Z}$ corresponding to state *x*, π_{x} is the stationary mass at state *x*, and $\hat{\pi}^{x}$ is the steady-state vector after we flip the control gene at state *x*.

A SSD control policy, .., is found via the following procedure. For every pair of states *x* and $\tilde{x}$, compute $\pi_{\mathcal{U}} = \sum_{y\in\mathcal{U}} \pi_{y}$, $\hat{\pi}^{x}_{\mathcal{U}} = \sum_{y\in\mathcal{U}} \hat{\pi}^{x}_{y}$ and $\hat{\pi}^{\tilde{x}}_{\mathcal{U}} = \sum_{y\in\mathcal{U}} \hat{\pi}^{\tilde{x}}_{y}$. If $\pi_{\mathcal{U}} \le \min(\hat{\pi}^{x}_{\mathcal{U}}, \hat{\pi}^{\tilde{x}}_{\mathcal{U}})$ then $\mu_{\text{SSD}}(x) = \mu_{\text{SSD}}(\tilde{x}) = 0$; otherwise, if $\hat{\pi}^{x}_{\mathcal{U}} \le \hat{\pi}^{\tilde{x}}_{\mathcal{U}}$, then $\mu_{\text{SSD}}(x)=1$ and $\mu_{\text{SSD}}(\tilde{x}) = 0$, and if $\hat{\pi}^{x}_{\mathcal{U}} \ge \hat{\pi}^{\tilde{x}}_{\mathcal{U}}$, then $\mu_{\text{SSD}}(x)=0$ and $\mu_{\text{SSD}}(\tilde{x}) = 1$.

When deriving the SSD control policy, all beneficial one-row perturbations are found independently. However, one can also iteratively find a state and corresponding control action having a one-row perturbation resulting in the largest reduction of undesirable mass. This is the *conservative steady-state distribution* (CSSD) control policy. We begin by assigning no control to all states. At each iteration, one selects a state $\underline{x}$ and corresponding control action at state $\underline{x}$ with a maximal reduction in undesirable mass. The algorithm runs iteratively until intervention at any additional state actually increases the stationary mass of undesirable states. One can theoretically guarantee that the located control policy will always induce undesirable stationary mass less than or equal to that of the uncontrolled network (Qian et al., 2009). By construction, μ_{CSSD} is a stationary and deterministic policy. Table 3 presents the algorithm, where π and $\mathcal{Z}$ denote the steady-state distribution and the fundamental matrix of the unperturbed network, respectively.

None of the control policies discussed thus far check if there is an introduction of mass at states representing complications or harmful side effects to healthy cells, even though these states are labeled as desirable. It is therefore prudent to restrict the shift of mass to only those states associated with known healthy phenotypes. To meet this new criterion, we partition $\mathcal{D}$ further into two categories of .., representing known healthy states, and $\mathcal{D}_{\text{A}}$, representing *ambiguous* states where we have little or no prior knowledge and thus may lead to complications or damages (Qian & Dougherty, 2012).

The goal is now to find a control policy that shifts the undesirable probability mass as much as possible and at the same time keeps the ambiguous probability mass of each state below certain level. The *constrained SSD (conSSD)* and *constrained CSSD (conCSSD)* control policies aim to find such (suboptimal) policies (Qian & Dougherty, 2012).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. CSSD control policy algorithm (Qian et al., 2009)

Set $\underline{\pi}=\pi$ and $\underline{\mathcal{Z}} = \mathcal{Z}$; Assign the set of states with no control to the set of all states: $\mathcal{O} = \mathcal{S}$;
Repeat
Δ=0;
For all states $x \in \mathcal{O}$ do
Compute $\Delta\pi_{\underline{u}} = \underline{\pi}_{\underline{u}} - \hat{\underline{\pi}}^{x}_{\underline{u}}$;
If $\Delta\pi_{\underline{u}} > \underline{\Delta}$ then
$\underline{\Delta} = \Delta\pi_{\underline{u}}$; Assign the best state to control: $\underline{x}$=x; $\underline{\pi} = \hat{\pi}^{x}$;
End
If $\underline{\Delta}$>0 then
$\mu_{\text{CSSD}}(x)$=1 and $\mu_{\text{CSSD}}(\tilde{\underline{x}}) = 0$; $\mathcal{O} = \mathcal{O} \setminus \{\underline{x}, \tilde{\underline{x}}\}$; Update $\underline{\mathcal{Z}}$ for the rank-one perturbed network accordingly (refer to (Qian et al., 2009) for the updating equation);
Until $\underline{\Delta}$=0;

We denote by $\mu_{\text{conSSD}} : \mathcal{S} \to \mathcal{A}$ and $\mu_{\text{conCSSD}} : \mathcal{S} \to \mathcal{A}$ the conSSD and conCSSD control policies, respectively. The algorithms presented in Table 4 and Table 5 are similar to those of SSD and CSSD, respectively, except that they take into account an upper-bound, $\delta \geq 0$, on the long-run probability mass of ambiguous states after control. This will represent the maximum amount of risk one is willing to accept with regard to introducing probability mass at an ambiguous state while implementing the intervention policy. In practice, expert medical knowledge would be employed to identify the set $\mathcal{D}_{\text{A}}$ and upper bound δ (Qian & Dougherty, 2012). As a consequence of adding this stringent criterion, it is expected that these constrained control policies become less effective in terms of shifting the undesirable mass.

Optimal Phenotype Alteration

We now find optimal intervention policies with respect to the objectives and constraints. As we will see, the problem of maximal phenotype alteration is a special case of constrained MDPs (Altman, 1999). Therefore, we begin with a brief review of MPDs, where we focus on the average cost objective function. In general, the decision maker minimizes one objective function subject to constraints on possibly one or more additional objective functions. The basic components of a constrained MDP with average cost are the same as those of discounted cost, but with additional cost functions. Therefore, we adopt the same notations, and only define a new immediate cost function as $r : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \Re^{D}$, which is a D-dimensional non-negative and bounded vector corresponding to D constraints. We also denote the TPM of the controlled chain by $\mathcal{P}(a)$, and assume that the chain is ergodic under any deterministic policy. For an initial state $x \in \mathcal{S}$ and a control policy $\mu = \{\mu_1, \mu_2, \ldots\} \in \mathcal{M}$, the infinite-horizon expected average costs are defined by

$$G(x,\mu) = \limsup_{N\to\infty} E^{\mu}_{x} \frac{1}{N+1} \sum_{t=0}^{N} g(Z_t, A_t, Z_{t+1}), \tag{37}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. conSSD control policy algorithm (Qian & Dougherty, 2012)

For all pairs of states $x, \tilde{x} \in \mathcal{O}$
Compute $\hat{\pi}^{x}$ and $\hat{\pi}^{\tilde{x}}$;
If $\max_{y \in \mathcal{D}_{\Lambda}} \hat{\pi}_{y}^{x} > \delta$ then
$\hat{\pi}_{\mathcal{U}}^{x} = 1.0$; // Make the control action not admissible
Else
$\hat{\pi}_{\mathcal{U}}^{x} = \sum_{y \in \mathcal{U}} \hat{\pi}_{y}^{x}$;
If $\max_{y \in \mathcal{D}_{\Lambda}} \hat{\pi}_{y}^{\tilde{x}} > \delta$ then
$\hat{\pi}_{\mathcal{U}}^{\tilde{x}} = 1.0$; // Make the control action not admissible
Else
$\hat{\pi}_{\mathcal{U}}^{\tilde{x}} = \sum_{y \in \mathcal{U}} \hat{\pi}_{y}^{\tilde{x}}$;
If $\pi_{\mathcal{U}} \leq \min(\hat{\pi}_{\mathcal{U}}^{x}, \hat{\pi}_{\mathcal{U}}^{\tilde{x}})$
$\mu_{\text{conSSD}}(x) = \mu_{\text{conSSD}}(\tilde{x}) = 0$;
Else
If $\hat{\pi}_{\mathcal{U}}^{x} \leq \hat{\pi}_{\mathcal{U}}^{\tilde{x}}$
$\mu_{\text{conSSD}}(x)$=1 and $\mu_{\text{conSSD}}(\tilde{x}) = 0$;
Else
$\mu_{\text{conSSD}}(x)$=0 and $\mu_{\text{conSSD}}(\tilde{x}) = 1$;
End

Table 5. conCSSD control policy algorithm (Qian & Dougherty, 2012)

Set $\underline{\pi}$=π and $\underline{\mathcal{Z}} = \mathcal{Z}$; Assign the set of states with no control to the set of all states: $\mathcal{O} = \mathcal{S}$;
Repeat
$\underline{\Delta}$=–0.1;
For all states $x \in \mathcal{O}$ do
Compute $\Delta\pi_{\mathcal{U}} = \underline{\pi}_{\mathcal{U}} - \underline{\hat{\pi}}_{\mathcal{U}}^{x}$;
If $\max_{y \in \mathcal{D}_{\Lambda}} \underline{\hat{\pi}}_{y}^{x} > \delta$ then
$\Delta\pi_{\mathcal{U}} = -0.1$; // Make the control action not admissible
If $\Delta\pi_{\mathcal{U}} > \underline{\Delta}$ then
$\underline{\Delta} = \Delta\pi_{\mathcal{U}}$; Assign the best state to control: $\underline{x}$=x; $\underline{\pi} = \hat{\pi}^{x}$;
End
If Δ>0 then
$\mu_{\text{conCSSD}}(\underline{x})$=1 and $\mu_{\text{conCSSD}}(\underline{\tilde{x}}) = 0$; $\mathcal{O} = \mathcal{O} \setminus \{\underline{x}, \underline{\tilde{x}}\}$; Update $\underline{\mathcal{Z}}$ for the rank-one perturbed network accordingly (refer to (Qian et al., 2009) for the updating equation);
Until $\underline{\Delta} \leq 0$;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$R_d(x,\mu) = \limsup_{N\to\infty} E_x^{\mu} \frac{1}{N+1} \sum_{t=0}^{N} r_d(Z_t, A_t, Z_{t+1}), \text{ for } d = 1,2,\ldots,D. \tag{38}$$

Given $C_1, C_2, \ldots, C_D \geq 0$, we can formulate the optimization problem in its most general form for any initial state $x \in \mathcal{S}$,

$$\begin{aligned} &\min_{\mu\in\mathcal{M}} G(x,\mu) \\ &\text{subject to } R_d(x,\mu) \leq C_d \text{ for all } d = 1,2,\ldots,D. \end{aligned} \tag{39}$$

All policies that satisfy the constraints are called feasible. Let $G^*(x)$ denote a solution of (39). A feasible policy μ^* is optimal if $G^*(x) = G(x,\mu^*)$, for all $x \in \mathcal{S}$. Although, the search space $\mathcal{M}$ is infinite (possibly uncountable), we consider a set $\mathcal{M}_S \subset \mathcal{M}$, of stationary policies without loss of optimality. $\mathcal{M}_S$ includes those policies for which μ_t is time invariant (not dependent on t) and only a function of z_t for any t. In other words, the probabilities that a control policy μ assigns to different actions only depends on the current state observed and not the time. Note that the set of stationary and deterministic policies, $\mathcal{M}_D$, is a subset of $\mathcal{M}_S$.

It can be shown that may be transformed into a LP with its dual being (Altman, 1999):

$$\begin{aligned} &\min_{\{\nu_{xa}\}} \sum_x \sum_a \nu_{xa} g(x,a) \\ &\text{subject to } \begin{cases} \sum_x \sum_a \nu_{xa} r_d(x,a) \leq C_d, \text{ for all } d = 1,2,\ldots,D \\ \sum_a \nu_{xa} = \sum_y \sum_a \nu_{ya} p_{yx}(a), \text{ for all } x \in \mathcal{S} \\ \sum_x \sum_a \nu_{xa} = 1 \\ \nu_{xa} \geq 0, \text{ for all } x \in \mathcal{S},\ a \in \mathcal{A} \end{cases} \end{aligned} \tag{40}$$

where $g(x,a) = \sum_y p_{xy}(a) g(x,a,y)$ and $r_d(x,a) = \sum_y p_{xy}(a) r_d(x,a,y)$ for each d=1,2,…,D. Letting $\left\{\nu_{xa}^*\right\}$ be a minimizing argument of (40), one can recover an optimal stationary policy $\mu^* \in \mathcal{M}_S$ by

$$\mu^*(a \mid x) = \frac{\nu_{xa}^*}{\sum_a \nu_{xa}^*}. \tag{41}$$

If D>0 and is feasible, then μ^* requires randomization (choosing between the actions in $\mathcal{A}$ through a randomized mechanism) in at most D states (Altman, 1999; Ross, 1989). However, if there are no constraints on the expected average cost criteria, i.e., D=0, then there is no randomization and $\mu^* \in \mathcal{M}_D$.

In our maximal phenotype alteration problem, we desire an intervention policy that optimally shifts the long-run probability mass of undesirable states towards desirable states, without introducing new mass at ambiguous states. This objective essentially concerns the long-run expected frequencies of observing undesirable states which is encoded in the variables $\{\nu_{xa}\}$ marginalized over the actions, i.e., $\left\{\sum_a \nu_{xa}\right\}$.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Let $\bar{\mu}(x) \equiv 0$ denote the stationary and deterministic control policy of *no-intervention*. For any policy $\mu \in \mathcal{M}_{\mathrm{S}}$, the amount of *shift* in the aggregated probability of undesirable states is defined as

$$\Delta\pi_{\mathcal{U}}(\mu) = \pi_{\mathcal{U}}(\bar{\mu}) - \pi_{\mathcal{U}}(\mu), \tag{42}$$

where $\pi_{\mathcal{U}}(\mu) = \sum_{x\in\mathcal{U}} \pi_x(\mu)$ and $\boldsymbol{\pi}(\boldsymbol{\mu})$ is the steady-state row vector of the controlled Markov chains governed under policy $\boldsymbol{\mu}$. In general, $-1 \le \Delta\pi_{\mathcal{U}}(\mu) \le 1$ and our goal is to maximize it subject to some constraints on the elements of $\boldsymbol{\pi}(\boldsymbol{\mu})$, which are usually in the form of $\boldsymbol{\pi}_x(\boldsymbol{\mu}) \le \boldsymbol{\gamma}_x$ for every $x \in \mathcal{D}_{\mathrm{A}}$ and constants $\boldsymbol{\gamma}_x \ge 0$.

First, note that maximizing $\Delta\pi_{\mathcal{U}}(\mu)$ is equivalent to minimizing $\pi_{\mathcal{U}}(\mu)$. Now, let the immediate cost functions g and r_d be defined as

$$g(x,a,y) = \begin{cases} 1, & \text{if } x \in \mathcal{U} \\ 0, & \text{otherwise} \end{cases} \tag{43}$$

$$r_d(x,a,y) = \begin{cases} 1, & \text{if } x \in \mathcal{D}_{\mathrm{A}} \\ 0, & \text{otherwise} \end{cases} \tag{44}$$

for any action $a \in \mathcal{A}$, any state $y \in \mathcal{S}$ and all d. Using these cost functions, we will count the number of visits to states in $\mathcal{U}$ and $\mathcal{D}_{\mathrm{A}}$. Thus, for a policy $\mu \in \mathcal{M}_{\mathrm{S}}$ and an initial state x, equations (37) and (38) provide in the steady-state mass of undesirable and ambiguous states, respectively. The constrained maximal phenotype alteration problem can thus be formulated as

$$\min_{\{\nu_{xa}\}} \sum_x \sum_a \nu_{xa}$$
$$\text{subject to } \begin{cases} \sum_a \nu_{xa} \le \gamma_x, \text{ for each } x \in \mathcal{D}_{\mathrm{A}} \\ \sum_a \nu_{xa} = \sum_y \sum_a \nu_{ya} p_{yx}(a), \text{ for all } x \in \mathcal{S} \\ \sum_x \sum_a \nu_{xa} = 1 \\ \nu_{xa} \ge 0, \text{ for all } x \in \mathcal{S},\ a \in \mathcal{A} \end{cases} \tag{45}$$

Once the solution $\left\{\nu^*_{xa}\right\}$ is found, one can recover the optimal stationary policy using (41). If there are no ambiguous states, then the optimal policy is stationary and deterministic. We refer to this policy as the *optimal unconstrained (UC)* policy and denote it by μ^*_{UC}. If there are $|\mathcal{D}_{\mathrm{A}}|$ ambiguous states then the optimal policy may not be deterministic; however, it belongs to $\mathcal{M}_{\mathrm{S}}$ and requires randomization in at most $|\mathcal{D}_{\mathrm{A}}|$ states. In this case, we call this policy *optimal phenotypically constrained (PC)* and denote it by μ^*_{PC}.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Simulation on Networks

We test our method on synthetically generated networks and a real network modeling metastatic melanoma.

Synthetic Networks

For phenotypically constrained problems and for any PBN, we define $\mathcal{D}_{\mathrm{A}} = \{x \in \mathcal{D} : \pi_x \leq 1/2^n\}$, where π is the steady-state vector of the uncontrolled network. We also set $\gamma_x = \max_{x \in \mathcal{D}_{\mathrm{A}}} \pi_x$ for all $x \in \mathcal{D}_{\mathrm{A}}$.

We synthetically generate 2500 instantaneously random PBNs with seven genes and report both the average performance across all networks as well as the empirical complementary CDF of

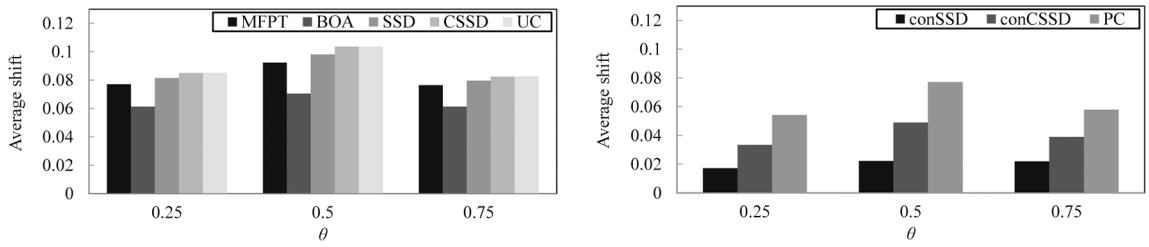
$$\Delta(\mu) = \Delta\pi_u(\mu^*) - \Delta\pi_u(\mu) = \pi_u(\mu) - \pi_u(\mu^*). \tag{46}$$

Here, $\Delta(\mu)$ quantifies the gain made by using an optimal policy, μ^*, over a suboptimal one, μ. Clearly, $0 \leq \Delta(\mu) \leq 1$. For each network, we calculate $\Delta(\mu)$ for all six suboptimal policies discussed earlier in this section and both optimal policies, for both the unconstrained and constrained problems, separately.

We assume that each PBN is generated from four equally likely constituent Boolean networks with the maximum number of predictors for each Boolean function set to three. We assume that the probability that each Boolean regulatory function takes on the value 1 is a random variable whose values are taken from a beta distribution with mean $\theta \in \{0.25, 0.5, 0.75\}$ and standard deviation 0.01. This quantity controls the dynamics of a randomly generated Boolean network. We also assume a random gene perturbation probability of 0.001. We define undesirable states, $\mathcal{U}$, to be those for which the most significant bit in their binary representation is zero, and assume that the control gene is the most significant bit of states.

Figure 8 shows the shift in the probability mass of undesirable states induced by the optimal and suboptimal policies, defined in (42), averaged across all PBNs for each values of θ. These graphs clearly illustrate the optimality of μ^*_{UC} and μ^*_{PC} compared to their suboptimal counterparts. Among the unconstrained suboptimal policies, μ_{CSSD} yields the best results. However, μ_{conCSSD} does not induce as much shift as μ^*_{PC}, which is due to the fact that the search space of the conCSSD policy belongs to the class $\mathcal{M}_{\mathrm{D}}$ of policies, whereas μ^*_{PC} belongs to the larger class $\mathcal{M}_{\mathrm{S}}$. It is worth noting that the amount of average shift induced by μ^*_{PC} is significantly less than μ^*_{UC}, which can be attributed to the additional constraints imposed on the optimization problem.

Figure 8. Average shift for different policies across 2500 random PBNs with seven genes: (Left) unconstrained policies; (Right) Phenotypically constrained policies



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

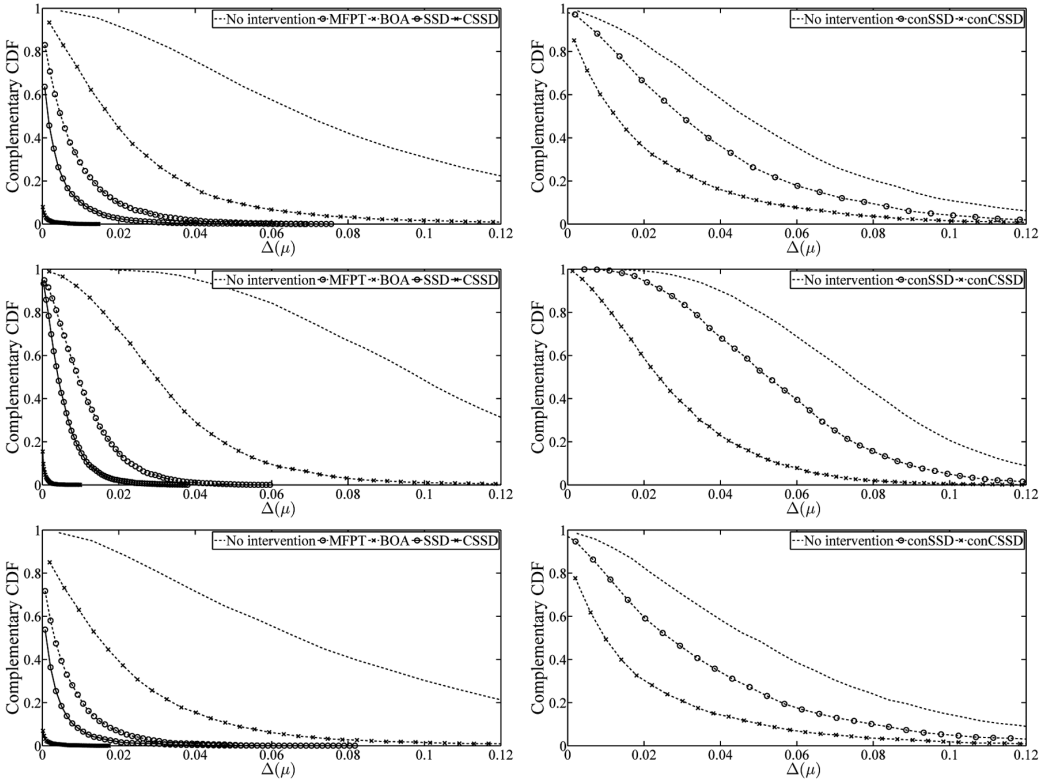
Figure 9 presents the empirical complementary CDF of $\Delta(\mu)$ for different suboptimal interventions. The graphs indicate that μ^*_{UC} and μ_{CSSD} induce similar amounts of shift with a high probability. It is also evident that μ^*_{PC} outperforms μ_{conCSSD} with a (realistically) significant probability; for instance, with probability 0.1 using μ^*_{PC} may yield an improvement of more than 5% over μ_{conCSSD}.

Real Network

It has been demonstrated that by manipulating the concentration level of the Wnt5a protein secreted by a melanoma cell line, one can directly affect the metastatic status of the cell as measured by standard *in vitro* assays for metastasis (Bittner et al., 2000; Dissanayake et al., 2008; Weeraratna et al., 2002), which suggests that reducing the WNT5A gene's action could reduce the risk of a melanoma metastasizing. Therefore, it is desirable to down-regulate WNT5A as much as possible through appropriate intervention mechanisms (Datta et al., 2003; Qian & Dougherty, 2008).

We use a Boolean network composed of ten nodes extracted from a set of 587 genes (Kim et al., 2002). The Boolean regulatory relationships between these genes/proteins are presented in Table 6. Note there is a single context when we construct a PBN from this network. We label states with an up-regulated WNT5A as undesirable. We set the gene perturbation probability to 0.001. We assume that up-

Figure 9. Empirical complementary CDF of $\Delta(\mu)$ for different intervention policies: (Top-left) θ=0.25 and μ^*_{UC}; (Top-right) θ=0.25 and μ^*_{PC}; (Middle-left) θ=0.5 and μ^*_{UC}; (Middle-right) θ=0.5 and μ^*_{PC}; (Bottom-left) θ=0.75 and μ^*_{UC}; (Bottom-right) θ=0.75 and μ^*_{PC}



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Boolean network of a metastatic melanoma network

Node Name	Gene Name	Predictor Function
v^1	WNT5A	$\left(\text{S100P } \textbf{AND } \text{MMP3 } \textbf{AND } \overline{\text{PHOC}}\right) \textbf{ OR } \left(\overline{\text{MMP3}} \textbf{ AND } \text{PHOC}\right)$
v^2	pirin	$\left(\overline{\text{WNT5A}} \textbf{ AND } \overline{\text{S100P}} \textbf{ AND } \text{MMP3}\right) \textbf{ OR } \left(\text{WNT5A } \textbf{AND } \overline{\text{S100P}} \textbf{ AND } \overline{\text{MMP3}}\right)$
v^3	S100P	MART1
v^4	RET1	$\left(\overline{\text{WNT5A}} \textbf{ AND } \text{pirin } \textbf{AND } \text{RET1}\right) \textbf{OR} \left(\overline{\text{pirin}} \textbf{ AND } \text{RET1}\right)$
v^5	MMP3	$\left(\text{RET1 } \textbf{AND } \text{synuclein}\right) \textbf{ OR } \overline{\text{synuclein}}$
v^6	PHOC	$\left(\overline{\text{RET1}} \textbf{ AND } \overline{\text{MART1}}\right) \textbf{ OR } \left(\text{RET1 } \textbf{AND } \text{MART1 } \textbf{AND } \text{STC2}\right)$
v^7	MART1	MART1
v^8	HADHB	$\left(\text{WNT5A } \textbf{AND } \text{MMP3}\right) \textbf{ OR } \left(\overline{\text{MMP3}} \textbf{ AND } \overline{\text{synuclein}}\right) \textbf{ OR } \left(\text{WNT5A } \textbf{AND } \overline{\text{MMP3}} \textbf{ AND } \text{synuclein}\right)$
v^9	synuclein	$\left(\overline{\text{WNT5A}} \textbf{ AND } \overline{\text{MART1}} \textbf{ AND } \overline{\text{STC2}}\right) \textbf{ OR } \left(\text{RET1 } \textbf{AND } \overline{\text{MART1}} \textbf{ AND } \text{STC2}\right) \textbf{ OR } \text{MART1}$
v^{10}	STC2	$\overline{\text{S100P}}$

regulation of WNT5A is undesirable, and set ambiguous states in $\mathcal{D}$ to non-attractor states of the Boolean network, so that the control policy will not introduce new attractors. For the control genes, we choose each gene in the network, except WNT5A, as the control gene, in turn.

The aggregated probability mass of undesirable states for the uncontrolled network is $\pi_{_U}(\overline{\mu}) = 0.1711$. Table 7 shows the shift made by all control policies based on different control genes. The maximum shift induced by μ^*_{UC} occurs when either MMP3 or PHOC are the control genes. However, the maximum shift for μ^*_{PC} occurs when MART1 is the control gene, which is an immunogen that the body sometimes recognizes and attacks via the immune system (Dissanayake et al., 2008; Qian & Dougherty, 2012).

CONCLUSION

Modeling tumor heterogeneity and the degree to which tumor cells respond to antitumor drugs are crucial to developing optimal cancer treatment. Using a recently proposed model for tumor growth inhibition, we have characterized uncertainties associated with drug response at different levels and built a probabilistic model for the entire intervention process. While accounting for practical effects arising from therapeutic constraints, we assumed PBNs as our model for GRNs and derived optimal control policies with random-length duration of treatment effectiveness. This was accomplished by taking a unified approach in the framework of finite-state MDPs with partially controlled decision intervals and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. Amount of shift in the metastatic melanoma network for different control genes and policies

Policy Gene	MFPT	BOA	SSD	CSSD	Optimal UC	conSSD	conCSSD	Optimal PC
pirin	0.1632	0.1630	0.1630	0.1632	0.1632	0.0003	0.0043	0.0548
S100P	0.1679	0.1673	0.1679	0.1679	0.1681	0.1548	0.0080	0.0139
RET1	0.1681	0.1681	0.1646	0.1681	0.1681	0.0000	0.0005	0.1473
MMP3	0.1701	0.1701	0.1701	0.1701	0.1701	0.0054	0.0059	0.0076
PHOC	0.1701	0.1701	0.1701	0.1701	0.1701	0.0041	0.0057	0.0076
MART1	0.1684	0.1684	0.1684	0.1684	0.1684	0.0165	0.0016	0.1682
HADHB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
synuclein	0.0029	0.0029	0.0029	0.0029	0.0029	0.0013	0.0029	0.0029
STC2	0.1671	0.1278	0.1674	0.1674	0.1674	0.1146	0.0040	0.0070

infinite-horizon discounted cost. This framework is general enough to include several other, simple but practical, intervention models. We presented a simulation study on randomly generated instantaneously random PBNs and demonstrated that optimal intervention policies outperform suboptimal ones with a significant probability. We also showed the effectiveness of optimal intervention policies on mutated mammalian cell cycle regulation.

Although the control policies derived from this procedure are optimal, their optimality is relative to subjective cost functions defined by consulting physicians or biologists. We resolved this issue by addressing optimality directly relative to the long-run behavior of the underlying network and phenotypes of interest. We utilized a linear programming approach to optimally move the steady-state mass of undesirable states toward desirable states. We provided a simulation study on synthetic PBNs with different dynamical properties and a metastatic melanoma network, where we demonstrated the superiority of optimal methods over several suboptimal approaches.

Besides the approaches presented here, many additional methods have been proposed: (1) to constrain the intervention so as to limit the expected number of interventions (Ching et al., 2009; Faryabi et al., 2008); (2) to deal with biological model uncertainties via robust intervention policies (Pal, Datta, & Dougherty, 2008, 2009; Yousefi & Dougherty, 2014); (3) to address unpredictable effects of a drug on multiple off-target genes (Yousefi & Ivanov, 2013); (4) to address asynchronicity of updating rules in gene networks (Faryabi et al., 2008); and (5) to mitigate the computational burden from complex gene networks (Akutsu, Hayashida, Ching, & Ng, 2007; Faryabi, Datta, & Dougherty, 2007; Ivanov, Simeonov, Ghaffari, Qian, & Dougherty, 2010; Ng, Zhang, Ching, & Akutsu, 2006; Vahedi et al., 2008).

Although the applicability of many of the discussed approaches may appear to hinge on first demonstrating the validity of PBNs as a model for GRNs, these approaches are not restricted to PBNs. In fact, the main assumption underlying most of these approaches is that the dynamics of GRNs can be well-approximated by Markovian stochastic processes. This assumption is not restrictive once we realize that the class of Markovian models encompasses a large body of both deterministic and stochastic dynamical models of GRNs (Kim et al., 2002; Goutsias & Jenkinson, 2013). As long as we can characterize the joint probability distribution of the stochastic process of interest and have a well-defined objective func-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tion, with reasonable assumptions on the intervention model, we can derive either an analytical or an approximate solution to various intervention problems. The question now is whether inferring the joint probability distribution is possible, both computationally and experimentally. To answer this question, we must first answer two even more fundamental questions: (1) what assumptions are we willing to impose on the form of this joint probability distribution? and (2) what types of data, and how much, do we need to collect to achieve an accuracy, in terms of model prediction error, within a desirable range? Nonetheless, it is this author's belief that methods for optimal or approximate intervention will play an important role in the near future when the medical community is ready to adopt model-based design of intervention strategies for cancer treatment.

REFERENCES

Akutsu, T., Hayashida, M., Ching, W.-K., & Ng, M. K. (2007). Control of Boolean networks: Hardness results and algorithms for tree structured networks. *Journal of Theoretical Biology*, *244*(4), 670–679. doi:10.1016/j.jtbi.2006.09.023 PMID:17069859

Altman, E. (1999). *Constrained Markov Decision Processes* (Vol. 7). London, U.K.: Chapman & Hall/CRC.

Bertsekas, D. P. (2005). *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific.

Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., ... Trent, J. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature*, *406*(6795), 536–540. doi:10.1038/35020115 PMID:10952317

Brun, M., Dougherty, E. R., & Shmulevich, I. (2005). Steady-state probabilities for attractors in probabilistic Boolean networks. *Signal Processing*, *85*(10), 1993–2013. doi:10.1016/j.sigpro.2005.02.016

Brunton, L. L., Chabner, B., & Knollmann, B. C. (2010). *Goodman & Gilman's The Pharmacological Basis of Therapeutics, 12e*. New York, NY: McGraw-Hill Professional.

Cao, Y., & Samuels, D. C. (2009). Discrete stochastic simulation methods for chemically reacting systems. *Methods in Enzymology*, *454*, 115–140. doi:10.1016/S0076-6879(08)03805-6 PMID:19216925

Chaouiya, C., Remy, E., & Thieffry, D. (2008). Petri net modelling of biological regulatory networks. *Journal of Discrete Algorithms*, *6*(2), 165–177. doi:10.1016/j.jda.2007.06.003

Chen, K.-C., Wang, T.-Y., Tseng, H.-H., Huang, C.-Y. F., & Kao, C.-Y. (2005). A stochastic differential equation model for quantifying transcriptional regulatory network in Saccharomyces cerevisiae. *Bioinformatics (Oxford, England)*, *21*(12), 2883–2890. doi:10.1093/bioinformatics/bti415 PMID:15802287

Ching, W. K., Zhang, S. Q., Jiao, Y., Akutsu, T., Tsing, N. K., & Wong, A. S. (2009). Optimal control policy for probabilistic Boolean networks with hard constraints. *IET Systems Biology*, *3*(2), 90–99. doi:10.1049/iet-syb.2008.0120 PMID:19292563

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cojocaru, L., & Agur, Z. (1992). A theoretical analysis of interval drug dosing for cell-cycle-phase-specific drugs. *Mathematical Biosciences*, *109*(1), 85–97. doi:10.1016/0025-5564(92)90053-Y PMID:1591451

Datta, A., Choudhary, A., Bittner, M. L., & Dougherty, E. R. (2003). External Control in Markovian Genetic Regulatory Networks. *Machine Learning*, *52*(1), 169–191. doi:10.1023/A:1023909812213

Datta, A., & Dougherty, E. R. (2006). *Introduction to Genomic Signal Processing with Control*. Boca Raton, FL: CRC. doi:10.1201/9781420006674

Dayneka, N. L., Garg, V., & Jusko, W. J. (1993). Comparison of four basic models of indirect pharmacodynamic responses. *Journal of Pharmacokinetics and Pharmacodynamics*, *21*(4), 457–478. doi:10.1007/BF01061691 PMID:8133465

De Jong, H. (2002). Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, *9*(1), 67–103. doi:10.1089/10665270252833208 PMID:11911796

de Vladar, H. P., & González, J. A. (2004). Dynamic response of cancer under the influence of immunological activity and therapy. *Journal of Theoretical Biology*, *227*(3), 335–348. doi:10.1016/j.jtbi.2003.11.012 PMID:15019501

Derman, C. (1970). *Finite State Markovian Decision Processes* (Vol. 107). Orlando, FL: Academic Press, Inc.

Dissanayake, S. K., Olkhanud, P. B., O'Connell, M. P., Carter, A., French, A. D., Camilli, T. C., ... Weeraratna, A. T. (2008). Wnt5A regulates expression of tumor-associated antigens in melanoma via changes in signal transducers and activators of transcription 3 phosphorylation. *Cancer Research*, *68*(24), 10205–10214. doi:10.1158/0008-5472.CAN-08-2149 PMID:19074888

Dougherty, E. R., & Datta, A. (2005). Genomic signal processing: Diagnosis and therapy. *IEEE Signal Processing Magazine*, *22*(1), 107–112. doi:10.1109/MSP.2005.1407722

Dougherty, E. R., Pal, R., Qian, X., Bittner, M. L., & Datta, A. (2010). Stationary and structural control in gene regulatory networks: Basic concepts. *International Journal of Systems Science*, *41*(1), 5–16. doi:10.1080/00207720903144560

Dougherty, E. R., & Shmulevich, I. (2012). On the Limitations of Biological Knowledge. *Current Genomics*, *13*(7), 574–587. doi:10.2174/138920212803251445 PMID:23633917

El Samad, H., Khammash, M., Petzold, L., & Gillespie, D. (2005). Stochastic modelling of gene regulatory networks. *International Journal of Robust and Nonlinear Control*, *15*(15), 691–711. doi:10.1002/rnc.1018

Faryabi, B., Chamberland, J.-F., Vahedi, G., Datta, A., & Dougherty, E. R. (2008). Optimal intervention in asynchronous genetic regulatory networks. *IEEE Journal of Selected Topics in Signal Processing*, *2*(3), 412–423. doi:10.1109/JSTSP.2008.923853

Faryabi, B., Datta, A., & Dougherty, E. R. (2007). On approximate stochastic control in genetic regulatory networks. *IET Systems Biology*, *1*(6), 361–368. doi:10.1049/iet-syb:20070015 PMID:18203582

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Faryabi, B., Vahedi, G., Chamberland, J.-F., Datta, A., & Dougherty, E. R. (2008). Optimal Constrained Stationary Intervention in Gene Regulatory Networks. *EURASIP Journal on Bioinformatics & Systems Biology*, (1): 620767. PMID:18516276

Faure, A., Naldi, A., Chaouiya, C., & Thieffry, D. (2006). Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics (Oxford, England)*, *22*(14), e124–e131. doi:10.1093/bioinformatics/btl210 PMID:16873462

Friedman, N. (2004). Inferring cellular networks using probabilistic graphical models. *Science Signaling*, *303*(5659), 799–805. PMID:14764868

Giacinti, C., & Giordano, A. (2006). RB and cell cycle progression. *Oncogene*, *25*(38), 5220–5227. doi:10.1038j.onc.1209615 PMID:16936740

Goutsias, J., & Jenkinson, G. (2013). Markovian dynamics on complex reaction networks. *Physics Reports*, *529*(2), 199–264. doi:10.1016/j.physrep.2013.03.004

Heppner, G. H. (1984). Tumor Heterogeneity. *Cancer Research*, *44*(6), 2259–2265. PMID:6372991

Himmelstein, K. J., & Bischoff, K. B. (1973). Mathematical representations of cancer chemotherapy effects. *Journal of Pharmacokinetics and Pharmacodynamics*, *1*(1), 51–68. doi:10.1007/BF01060027

Huang, Y., Tienda-Luna, I., & Wang, Y. (2009). Reverse engineering gene regulatory networks. *IEEE Signal Processing Magazine*, *26*(1), 76–97. doi:10.1109/MSP.2008.930647 PMID:20046885

Iglesias, P. A., & Ingalls, B. P. (2010). *Control Theory and Systems Biology*. MIT Press.

Ivanov, I., Simeonov, P., Ghaffari, N., Qian, X., & Dougherty, E. R. (2010). Selection policy-induced reduction mappings for Boolean networks. *IEEE Transactions on Signal Processing*, *58*(9), 4871–4882. doi:10.1109/TSP.2010.2050314

Julius, A. A., Halasz, A., Sakar, M. S., Rubin, H., Kumar, V., & Pappas, G. J. (2008). Stochastic Modeling and Control of Biological Systems: The Lactose Regulation System of Escherichia Coli. *IEEE Transactions on Automatic Control*, *53*(Special Issue), 51–65. doi:10.1109/TAC.2007.911346

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, *22*(3), 437–467. doi:10.1016/0022-5193(69)90015-0 PMID:5803332

Kemény, J. G., & Snell, J. L. (1960). *Finite Markov Chains*. Princeton, NJ: Van Nostrand.

Kim, S., Li, H., Dougherty, E. R., Cao, N., Chen, Y., Bittner, M., & Suh, E. B. (2002). Can Markov chain models mimic biological regulation? *Journal of Biological System*, *10*(04), 337–357. doi:10.1142/S0218339002000676

Kyoda, K., Baba, K., Onami, S., & Kitano, H. (2004). DBRF–MEGN method: An algorithm for deducing minimum equivalent gene networks from large-scale gene expression profiles of gene deletion mutants. *Bioinformatics (Oxford, England)*, *20*(16), 2662–2675. doi:10.1093/bioinformatics/bth306 PMID:15166016

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Magni, P., Germani, M., De Nicolao, G., Bianchini, G., Simeoni, M., Poggesi, I., & Rocchetti, M. (2008). A Minimal Model of Tumor Growth Inhibition. *IEEE Transactions on Bio-Medical Engineering*, *55*(12), 2683–2690. doi:10.1109/TBME.2008.913420 PMID:19126447

Magni, P., Simeoni, M., Poggesi, I., Rocchetti, M., & De Nicolao, G. (2006). A mathematical model to study the effects of drugs administration on tumor growth dynamics. *Mathematical Biosciences*, *200*(2), 127–151. doi:10.1016/j.mbs.2005.12.028 PMID:16516246

Matsuno, H., Doi, A., Nagasaki, M., & Miyano, S. (2000). Hybrid Petri net representation of gene regulatory network. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, (pp. 341–352).

Ng, M., Zhang, S.-Q., Ching, W.-K., & Akutsu, T. (2006). A Control Model for Markovian Genetic Regulatory Networks. *Transactions on Computational Systems Biology V*, *4070*, 36–48. doi:10.1007/11790105_4

Pal, R., Datta, A., & Dougherty, E. R. (2006). Optimal infinite-horizon control for probabilistic Boolean networks. *IEEE Transactions on Signal Processing*, *54*(6), 2375–2387. doi:10.1109/TSP.2006.873740

Pal, R., Datta, A., & Dougherty, E. R. (2008). Robust intervention in probabilistic Boolean networks. *IEEE Transactions on Signal Processing*, *56*(3), 1280–1294. doi:10.1109/TSP.2007.908964

Pal, R., Datta, A., & Dougherty, E. R. (2009). Bayesian Robustness in the Control of Gene Regulatory Networks. *IEEE Transactions on Signal Processing*, *57*(9), 3667–3678. doi:10.1109/TSP.2009.2022872

Pe'er, D., & Hacohen, N. (2011). Principles and Strategies for Developing Network Models in Cancer. *Cell*, *144*(6), 864–873. doi:10.1016/j.cell.2011.03.001 PMID:21414479

Qian, X., & Dougherty, E. R. (2008). Effect of Function Perturbation on the Steady-State Distribution of Genetic Regulatory Networks: Optimal Structural Intervention. *IEEE Transactions on Signal Processing*, *56*(10), 4966–4976. doi:10.1109/TSP.2008.928089

Qian, X., & Dougherty, E. R. (2012). Intervention in Gene Regulatory Networks via Phenotypically Constrained Control Policies Based on Long-Run Behavior. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *9*(1), 123–136. doi:10.1109/TCBB.2011.107 PMID:21788674

Qian, X., Ivanov, I., Ghaffari, N., & Dougherty, E. R. (2009). Intervention in gene regulatory networks via greedy control policies based on long-run behavior. *BMC Systems Biology*, *3*(1), 61. doi:10.1186/1752-0509-3-61 PMID:19527511

Queinnec, I., Garcia, G., Niculescu, S.-I., & Tarbouriech, S. (2007). *Biology and control theory: current challenges* (Vol. 357). Springer. doi:10.1007/978-3-540-71988-5

Ross, K. W. (1989). Randomized and past-dependent policies for Markov decision processes with multiple constraints. *Operations Research*, *37*(3), 474–477. doi:10.1287/opre.37.3.474

Shmulevich, I., & Aitchison, J. D. (2009). Deterministic and stochastic models of genetic regulatory networks. *Methods in Enzymology*, *467*, 335–356. doi:10.1016/S0076-6879(09)67013-0 PMID:19897099

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shmulevich, I., Dougherty, E. R., Kim, S., & Zhang, W. (2002). Probabilistic Boolean networks: A rule-based uncertainty model for gene regulatory networks. *Bioinformatics (Oxford, England)*, *18*(2), 261–274. doi:10.1093/bioinformatics/18.2.261 PMID:11847074

Shmulevich, I., Dougherty, E. R., & Zhang, W. (2002). Control of stationary behavior in probabilistic Boolean networks by means of structural intervention. *Journal of Biological System*, *10*(04), 431–445. doi:10.1142/S0218339002000706

Shmulevich, I., Dougherty, E. R., & Zhang, W. (2002). Gene perturbation and intervention in probabilistic Boolean networks. *Bioinformatics (Oxford, England)*, *18*(10), 1319–1331. doi:10.1093/bioinformatics/18.10.1319 PMID:12376376

Simon, R., & Norton, L. (2006). The Norton-Simon hypothesis: Designing more effective and less toxic chemotherapeutic regimens. *Nature Clinical Practice. Oncology*, *3*(8), 406–407. doi:10.1038/ncponc0560 PMID:16894366

Vahedi, G., Faryabi, B., Chamberland, J.-F., Datta, A., & Dougherty, E. R. (2008). Intervention in gene regulatory networks via a stationary mean-first-passage-time control policy. *IEEE Transactions on Bio-Medical Engineering*, *55*(10), 2319–2331. doi:10.1109/TBME.2008.925677 PMID:18838357

Vahedi, G., Faryabi, B., Chamberland, J.-F., Datta, A., & Dougherty, E. R. (2009). Optimal intervention strategies for cyclic therapeutic methods. *IEEE Transactions on Bio-Medical Engineering*, *56*(2), 281–291. doi:10.1109/TBME.2008.2003092 PMID:19272947

Weeraratna, A. T., Jiang, Y., Hostetter, G., Rosenblatt, K., Duray, P., Bittner, M., & Trent, J. M. (2002). Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell*, *1*(3), 279–288. doi:10.1016/S1535-6108(02)00045-4 PMID:12086864

White, D. J. (1978). *Finite Dynamic Programming: An Approach to Finite Markov Decision Processes*. New York, NY: John Wiley & Sons, Inc.

Xiao, Y., & Dougherty, E. R. (2007). The impact of function perturbations in Boolean networks. *Bioinformatics (Oxford, England)*, *23*(10), 1265–1273. doi:10.1093/bioinformatics/btm093 PMID:17379691

Yousefi, M. R., Datta, A., & Dougherty, E. R. (2012). Optimal Intervention Strategies for Therapeutic Methods With Fixed-Length Duration of Drug Effectiveness. *IEEE Transactions on Signal Processing*, *60*(9), 4930–4944. doi:10.1109/TSP.2012.2202114

Yousefi, M. R., Datta, A., & Dougherty, E. R. (2013). Optimal Intervention in Markovian Gene Regulatory Networks With Random-Length Therapeutic Response to Antitumor Drug. *IEEE Transactions on Bio-Medical Engineering*, *60*(12), 3542–3552. doi:10.1109/TBME.2013.2272891 PMID:23864151

Yousefi, M. R., & Dougherty, E. R. (2013). Intervention in gene regulatory networks with maximal phenotype alteration. *Bioinformatics (Oxford, England)*, *29*(14), 1758–1767. doi:10.1093/bioinformatics/btt242 PMID:23630177

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yousefi, M. R., & Dougherty, E. R. (2014). A comparison study of optimal and suboptimal intervention policies for gene regulatory networks in the presence of uncertainty. *EURASIP Journal on Bioinformatics & Systems Biology*, (6): 1–13. PMID:24708650

Yousefi, M. R., & Ivanov, I. (2013). Optimal control of gene regulatory networks with uncertain intervention effects. In *Global Conference on Signal and Information Processing (GlobalSIP)*, (pp. 105–108). IEEE. 10.1109/GlobalSIP.2013.6736826

Zañudo, J. G. T., & Albert, R. (2015). Cell Fate Reprogramming by Control of Intracellular Network Dynamics. *PLoS Computational Biology*, *11*(4), e1004193. doi:10.1371/journal.pcbi.1004193 PMID:25849586

This research was previously published in Emerging Research in the Analysis and Modeling of Gene Regulatory Networks edited by Ivan V. Ivanov, Xiaoning Qian, and Ranadip Pal; pages 161-208; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 58
Human–Data Interaction in Healthcare

Federico Cabitza
Università degli Studi di Milano-Bicocca, Italy

Angela Locoro
Università degli Studi di Milano-Bicocca, Italy

ABSTRACT

In this chapter, we focus on an emerging strand of IT-oriented research, namely Human-Data Interaction (HDI) and on how this can be applied to healthcare. HDI regards both how humans create and use data by means of interactive systems, which can both assist and constrain them and the operational level of data work, which is both work on data and by data. Healthcare is a challenging arena where to test the potential of HDI towards a new, user-centered perspective on how to support and assess "data work". This is especially true in current times where data are becoming increasingly big and many tools are available for the lay people, including doctors and nurses, to interact with health-related data. This chapter is a contribution in the direction of considering health-related data through the lens of HDI, and of framing data visualization tools in this strand of research. The intended aim is to let the subtler peculiarities among different kind of data and of their use emerge and be addressed adequately. Our point is that doing so can promote the design of more usable tools that can support data work from a user-centered and data quality perspective and the evidence-based validation of these tools.

INTRODUCTION

Twenty-five years ago, medical informatics was defined as "dealing with the storage, retrieval and optimal use of biomedical data" (Shortliffe et al. 1990). At that time, little emphasis was put on the practices of data production, that is on how medical practice, and single stories of illness, care and recovery are represented, accounted and "datafied" in some valid and reliable manner. However, these practices, which include policies, rules, habits, conventions, tools and techniques, have always been intertwined with and affected by the available ITs, as well as by the expectations of the stakeholders on how to make sense

DOI: 10.4018/978-1-7998-1204-3.ch058

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and use of health-related data. Different perspectives on these expectations, and on what valuable health data are, lead to manifest chasms between primary use and other uses of health information, as often discussed in the specialist literature (Fitzpatrick and Ellingsen, 2013). To try to cross these chasms, we need to create the suitable language to describe the differences and give some operational definitions.

We distinguish between three different macro-types of data and the related processes in which these data are either produced, processed or consumed: namely, primary data, which come from a broad range of sources and are produced within a caring process to make its unfolding seamless and smooth (Berg, 1999); secondary data that are derived from the primary data for purposes different than care, like accounting and medical billing (Abdelhak, Grostick, & Hanken, 2012); and tertiary data, that are produced from the secondary data for any unanticipated need of the potential consumers of health services (see Figure 1).

To illustrate this tripartition, an analogy from the agriculture domain can be drawn (Locoro, 2016): primary data are like the produce of the land, which farmers grow for themselves as well as the external market. Secondary data are the product of a transformation of these primary data, like the one performed in food industry where vegetables are cleansed and chopped. Tertiary data are further transformed from secondary data to make them more easily consumable, that is suitable for and conveyed to a broader population of consumers in terms of information services, like fresh-cut vegetable products can be seen as the service to have vegetables already ready-to-eat.

The definitions mentioned above shed light on the relationship between data and their uses, which cannot be overstated. The tripartition that we propose reflects the different uses and practices in which data are produced and consumed and it calls for a specific area of research focusing on how to support people in interacting with their data of concern and in gaining insight from them: Human-Data Interaction (HDI).

BACKGROUND

As rightly noted by Hornung et al. (2015), the expression HDI is neither specific nor new. On the one hand, humans have always interacted with data, that is they have inscribed, read and communicated data in terms of simple marks, more complex symbols and any kind of meaningful signs[1]. On the other hand, this interaction has predominantly been mono-directional until interactive systems have allowed users to get access to data according to their peculiar and situated needs. In the computing literature, the expression has been used since the 1990s, e.g., by Kennedy et al. (1996), to address the needs of the users in the context of data intensive settings and hence "the problem of delivering personalized, context-aware, and understandable data from big datasets" (Cafaro, 2012). More recently a number of authors have tried to define Human-Data Interaction, by focusing on different aspects: Elmqvist (2011) and Cafaro (2012) focused on the material and embodied aspects of interaction; Crabtree and Mortier (2015) prefer to assimilate HDI to "the management and use of personal data in society at large" where *personal* here means that data are either "about individuals" or are delivered in terms of "personalized experiences". In particular, Elmqvist (2011) defines HDI as "the human manipulation, analysis, and sense making of large, unstructured, and complex datasets" and Cafaro sees HDI systems "as technologies that use embodied interaction to facilitate the users' exploration of rich datasets". We agree with Hornung et al. (2015) and Crabtree and Mortier (2014) in regard to the fact that "HDI does not only refer to embodied interaction, but to all kinds of interaction" and that "HDI should investigate 'data that affects people' ".

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

These reflections highlight the twofold meaning of HDI that we clarify in what follows. HDI is both a phenomenon and a research field that focuses on this phenomenon. As a phenomenon that is an object of research, we propose to see HDI as the class of actions where, on the one hand data are produced, processed and exploited by humans; this encompasses, at the two extremes of the action spectrum, both the "datafication" of facts, that is the process by which portions of the reality of interest are translated into the domain of words, symbols and numbers (through coding, classification, and measure[2]); and what we call data telling, that is the creation of accounts and stories that human can tell according to the data they make (a) sense of. On the other hand, HDI also regards actions in which humans are affected by data, that is in which their interpretations, opinions, beliefs and decisions are made by, or better yet according to, data.

Moreover, as a research area focusing on this kind of interaction it lies at the intersection of data science, data visualization and human-computer interaction. Researchers involved in HDI study how humans record and use data by means of Information Technologies; and design interactive systems that allow humans to retrieve and explore complex data sets in order to gain value in their learning, insight in their decisions, and feedback in their action. Considering HDI a distinct research field, although deeply grounded in the intersection between data visualization and human-computer interaction means to highlight the typical tenets of HCI about the centeredness of users in successful design and the importance to consider "complex contextual factors including the systems of beliefs, values and norms of the involved people" (Hornung et al. 2015) when designing systems by which humans can explore and interpret data according to their skills, competences, desires and needs. Moreover, it also means to borrow from HCI the attitude to evaluate the impact of these systems on the people life and investigate systematically "data production, collection, processing and use [by] focusing on the social impact they provoke" (Hornung et al. 2015).

In this chapter, we mainly focus on HDI as an evidence-based approach to the design of ICT systems. Our main contribution is a tripartite conceptual framework to help practitioners separate the concerns and address design related aspects peculiar to primary and tertiary data type, data work and data users. Three studies are also presented to exemplify the typical approaches within HDI and the related framework to the design, development and evaluation of healthcare supporting technologies. In particular, the first case refers to the heuristic evaluation of data visualization prioritization, the second one to a user test on data customization for citizens, and the third case to a research through design of a data visualization tool.

INTERACTION AS KEY TO UNDERSTAND HEALTH DATA

In this section we discuss how to consider interaction central to characterize and distinguish among the roles of health data, as an alternative to top-down, and abstract ways to classify them: in short we propose an analytical framework where considering who is going to use health data, and why (that is for what objective in what task) is more important than investigating the intrinsic attributes of data to understand how to better manage them (Strong, et al. 1997).

In virtue of these primary concerns, we distinguish between primary, secondary and tertiary (health) data. For a straight metonymy, we intend primary data as entangled with the primary use of data, by definition; secondary data with secondary use(s); tertiary data with tertiary uses. The "primary use" of health information is "to use it to directly support patient care", both by aiding medical decision-making and by ensuring continuity of care by all providers, that is both interpretation of medical signs (represented

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

by data) for decision making, and coordination among the actors involved "around" the patient (Berg, 1999). Secondary use regards both other uses of the same data collected for the primary use within the administrative domain, and the generation of derivative data for other aims than care, like billing and reimbursement, performance and care quality evaluation, resource planning and management, service design and public policy making. Tertiary use regards the heterogeneous uses through which tertiary data are put to the test of users' life: this can encompass the publication and dissemination of valuable indications for the citizens and the taxpayers about the available healthcare services, so as to orient them in the choice of the best healthcare provider (or just the closest one to their shelters – cf. Cabitza et al. 2015b); the policy makers and facility or agency manager, to facilitate the monitoring of suitable outcome and performance indicators and to enable the benchmarking and comparison of care facilities; as well as the researchers and epidemiologists to make sense of the great numbers of "similar" patients that are treated and detect trends and patterns of treatment appropriateness and efficacy.

On a more general level, three concepts characterize uses and data in each of these three ambits: primary use is related to what has been often called "data work" (Bossen et al. 2016). Secondary uses with data storing and processing. Tertiary uses with data value. Data storing and processing are self-explanatory concepts in computer science and regard the set-up of efficient data bases and the automation of accurate procedures by which to structure data, run searches and extract complex reports for disparate purposes. As secondary data and secondary processes usually entail little interaction by humans and these are always IT specialists and consultants who have always considered usability a relative concern, we will focus on the primary and secondary domains, and hence on data work and data value.

Data work is quite never defined in the specialist literature; although it is often mentioned, it is a slippery notion. We see it as a modern (but not at all recent) coinage after the word paperwork, of which it represents an abstraction, with respect to the medium of data representation; but also an extension, with respect to what people manage as data of their interest (besides accounting and resource management). As such data work is not only "working on data", typically producing new data in accounting for and recording a faithful representation of the work done; but also that portion of work that can be accomplished only by relying on some accurate data, i.e., "working by data". These two kinds of work are usually so deeply intertwined that distinguishing between them is useless and probably wrong: in the healthcare domain, the studies by Berg (Berg, 1999), for instance, clearly show that clinicians record data on the patient record not only to accumulate data for archival reasons (and for many other secondary uses), but also to coordinate with each other, articulate the resources around a medical case, and take informed decision in a written, distributed communication with themselves and the other colleagues taking care of the same patient. In healthcare, data work regards the additional effort paid by caregivers in making the record a "working record" (Fitzpatrick 2004), that is a resource capable of keeping disparate competencies and roles bound together and connected around the same cases over time and space.

On the other hand, tertiary use of tertiary data regards the concept of data value, that is how to exploit data to achieve some relevant goal for the user (even just to be informed about something), so that we can say that the user finds value in the data (also: she finds data valuable), or she gains value to some aim. Whether value is already in the data or rather it is created by an active stance of its consumers looks like an idle question: in a HDI perspective, value is the result of interacting with data and being capable to exploit them by tertiary users, that is lay people moved by unpredictable motives and toward unanticipated aims. In other words, on one hand data have got value if they are true and have been made accessible and comprehensible; on the other hand, their value lies in the comprehension itself, in the acquisition of true information, in learning notions, techniques, practices, and in the resulting knowledgeable behav-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

iors, in their turn producing some positive effect, on either the single person or her community[3]. Thus, data value can be defined as the potential of data (for its intrinsic qualities) to enable processes of value creation by someone for anyone (either herself, or the others) in some given context.

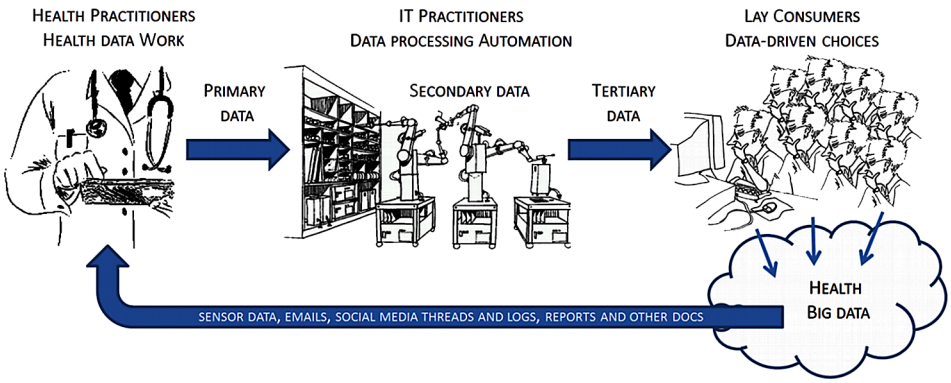
HDI CONTRIBUTIONS FOR THE QUALITY OF INFORMATION

While this tripartition is general enough to be applied in any domain where many different roles partake to the generation and transformation of data in the same supply chain, so to say, from the practitioner working in the field, to the final consumer of information services, its application to healthcare can be justified for its role in addressing the problem of the still relatively low quality of the data of patient records (Liaw et al. 2013), both in regard to completeness, accuracy and timeliness (Cabitza & Batini, 2016). Indeed it allows, on one hand, to highlight differences, separate concerns, and identifying ambits of use that are irreducible to one another; on the other hand, it allows to see the (existing) "seams" between these ambits (Berg and Goorman 1999) and therefore to address cooperation and interoperability without reckless neglect.

In regard to this latter point, it is time to acknowledge that chasms between primary and other uses exist and they go beyond the usual tension between clinical vs. administrative purposes because it is the concept of "adequate quality" that should be adopted to assess care-oriented data; likewise, crossing the chasm between the secondary and tertiary domains requires more than merely making secondary data more open and accessible, because also the end consumers' readiness to access, comprehend and exploit them, as well as their unpredictable purposes, are to be considered. Thus, the challenge to bridge the chasms mentioned above can be seen in terms of the problem of reusing primary data, which is "always entangled with the context of its production" (Berg and Goorman, 1999), in different contexts (either secondary or tertiary ones).

As Berg and Goorman note, reuse is possible only if data are made "transportable", that is sufficiently disentangled from their context of production; and this can occur only if specialized additional work is performed on data. So it is not a matter of improving, automating or changing current data work, but rather to invest on new and different data work, and corresponding new organizational roles (e.g., data nurses).

Figure 1. The three kinds of data and ambits proposed for HDI in healthcare



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In regard to the separation of concerns, our distinction between primary, secondary and tertiary data is quite different from other frameworks discussed in the Data Quality literature: for example, the one inspired by the manufacturing domain (Shankaranarayan et al., 2003) which distinguishes between raw data, component data items, that is semi-processed information and information products, which are composed out of these items. Primary data are not necessarily raw (Gitelman, 2013), because they are meaningful to and usable by the primary consumers that generated them. On the other hand, secondary and tertiary data are unfinished information products: the former ones are resources for specialist work (both clerical and managerial) within specific organizational boundaries and processes, while tertiary data result from the enactment of information services conceived for the non specialist and the external consumers, that is for the public. The distinction introduced by Shankaranarayan and his colleagues focuses on an incremental definition of the information product, and is often mentioned in regard to how Data Quality (DQ) can be assessed, monitored and continuously improved (e.g., Wang, 1998)

Differently from the Information Product perspective, our tripartite perspective focuses on the different roles that produce and consume different information products that can be considered "definitive" only in relation to their context of use. Consequently, a HDI approach to DQ issues grounds on the interpretation of quality of data as their "fitness for use" (Wang & Strong, 1996). Our proposal is aimed at detecting and taking the differences from the various uses in healthcare seriously (Bossen et al. 2016).

According to our metaphor, secondary data and secondary data-centered processes are like industrial processes, that need standardization to achieve regularity and efficiency and that involve relatively few people with specific skills and motives, like the competent and trained operators of assembly lines at the shop floor of plants. This is why HDI regards secondary data and processes only marginally, and to a much greater extent the primary and tertiary contexts. In primary settings, HDI addresses the challenges of fitting "data work" to work, that is of reconciling the articulation of cooperative activities performed by knowledgeable experts with their effort in continuous recording relevant events and information about what has been carried out, and often also of why and how. On the other hand, in the broader ambits of tertiary use, HDI deals with the manifold challenge to design data structures, visualization controls and interaction affordances, for both unintended uses and unanticipated needs, for both the practitioners without particular e-literacy and numeracy, and lay people from the general public.

We would argue that quality of information should not be assessed irrespective of the distinction between primary / secondary / tertiary uses, that is, by adopting the same metrics and methods in a context-independent manner: on the contrary, data used in care processes should be evaluated on the basis of the efficacy they enable appropriate and timely action (fit to use), also on the basis of work conventions and tacit knowledge that are difficult to bring back to the usual dimensions of accuracy, completeness and consistency (Gregory et al., 1995). This is in line with the idea to measure the quality of primary (health) data in terms of appropriateness and as a function of the health outcome they help to achieve. Bad quality should not be related to completeness, timeliness or, even, accuracy, since health practitioners can perfectly make sense of "bad" data and to some extent they even expect them (Bossen et al. 2016); but rather to "information failure" and its impact, that is the number of times data work is either direct or indirect cause of adverse event or near misses (Pipino & Lee, 2011). Despite the authorities that advocated for more research on this concept (Pipino & Lee, 201; Cabitza & Batini 2016), information failure (or error) is still under-researched and a taxonomy of cases to guide analysis and comparison still missing, for instance one that could distinguish between cases in which wrong or late decisions were carried out with accurate and timely data, respectively, and the other way round, a right and timely interpretation has been based on wrong and obsolete data, and so on[4].

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Similarly, further research is needed for the evaluation of the quality of tertiary data: this is highly correlated with the quality of the information services conceived for the general public, which should be evaluated in terms of the extent to which data are "informative" and can be appropriated by consumers and appreciated in their lives (i.e., according to its social value), that is if they affect decisions for the better.

This is where HDI can fit in and be a novel way to address long-standing problems of low quality that neither incentivizing structures nor sanctions seem to solve completely. Indeed, HDI covers three phases: design, development and evaluation of the systems by which to extract information and support knowledge in data-intensive application domains. As HDI approaches operate at the upper layer of the well known ISO-OSI standard, i.e., the application layer, or, better yet, the data visualization layer, the scope of HDI is not to tackle which data are to be shown in terms of ethical or security issues (which are tasks usually delegated to the lower layers of the ISO-OSI pillar, such as model and control layers, e.g., a database, a middleware, and so on), but how data are to be shown instead, e.g., in terms of ergonomic and usability issues. The scope of an HDI approach is mainly that of the design, development and evaluation of data visualization interfaces. In particular, HDI regards the user-centered elicitation of better requirements of configuration, adaptation and appropriation of big data analytics cockpits and dashboards to optimize usability and the user experience, i.e., efficiency, effectiveness and satisfaction; the application of End-User Development techniques and tools to allow end-users tweak the tools by which data are extracted and visualized; and user-centered methodology for the assessment and continuous improvement of the quality of the interaction of the humans with their data of interest, so as to reduce both information overloading and information funneling/complacency (Parasuraman and Manzey, 2010), and improve awareness. This also includes the exploration of new techniques toward better interactive visualization environments and above all better data-telling, that is the capability to build and share stories that can explain data and facilitate correct interpretations (e.g., in medical domain by adopting a natural frequency approach - Hoffrage et al., 2002) and to allow, instead of curbing, the social exchange within a community of data-users of multi-perspective, sound and viable interpretations around the data that are supplied by the computational systems to inform decision making and knowledgeable action.

Summarizing, HDI can contribute in addressing the reconciliation of needs in health data production and use by providing methods and techniques that investigate the following two research strands:

1. The design of interfaces that could promote mutual awareness between data producers and consumers: on the one hand, by increasing commitment and awareness of consequences in the producers of the primary data; and, at the same time, by raising awareness in the secondary consumers of the contextual and social nature of primary data and hence of their limitations. This can be achieved, for example, by endowing the interfaces by which data are collected and presented with specific affordances that adapt to the context according to specific business rules in order to convey the so called "awareness promoting information" (Cabitza and Simone, 2012); this can be done also by means of simple visual clues like text highlighting or side messages, which do not impose any behavior to the data producers but help their interpretation.
2. The design of interactive visual interfaces that could support the understandability of data analytics. The transformation of data into information services does not necessarily require a massive processing of data but rather the application of state-of-the-art human interaction techniques to develop interactive infographics and highly tailorable dashboards that enable user-friendly online analytical processing and hence the transformation, even by end users (Lieberman, Paternò, Klann, & Wulf, 2006), of secondary data into socially valuable information.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

AN HDI APPROACH TO FOSTER EVIDENCE-BASED DESIGN

The phrase evidence-based design is not totally new but still not really common, especially in the computer science literature and there is little wonder that also recently (Ammenwerth & Rigby, 2016) it has been applied to the fields of informatics that are closer to medicine and healthcare. In fact, the parallel with the approach to medical practice known as "Evidence-Based Medicine" (EBM)(VV.AA., 1992) is not coincidental, and indeed, it constitutes a sort of natural extension of that methodological empirical attitude to the design of computerized tools that can have an impact on medical decision and care quality. Shortly speaking, an evidence-based approach to IT design requires to back up any claim of utility or usability of a particular solution or tool with empirical studies that collect data from real users (including the perceptions of domain experts) and undertake standard statistical tests and group comparisons (Longhurst et al. 2013). Even more than this, an evidence-based approach must consider findings from the field of work in combination with the expertise of IT practitioners and the needs and wishes of the intended prospective users (e.g., patients and shopfloor doctors and nurses). Especially this latter aspect is typical in the Human-Computer Interaction field, and we also proposed it for studies accomplished within the HDI field (Cabitza et al., 2016).

To give more concreteness to the notion of HDI and evidence-based design, in the rest of this chapter we will report about three small-scale user studies in the domain of health information. The common element of these studies is the collection of empirical "proofs" (that is evidence) about solutions that were designed to increase the perceived value of interfaces and data visualization tools in the healthcare sector. To our knowledge the most relevant of the similar initiatives is the "Visualizing Health" project (http://www.vizhealth.org/), developed by the University of Michigan and the Robert Wood Johnson Foundation. In this project, which is subtitled "a scientifically vetted style guide for communicating health data", the involved researchers compared tens of different styles by which to communicate risk-related information in order to select the best ones for the general public according to a series of empirical tests involving real users. The main finding, which we also back up, is that when it comes to presenting health information, there is no single "best" graphic, as this depends on the intended goal and, in a subtler way, on what matters most to the intended users of the diagrams and infographics (Cabitza et al. 2015b). This calls for an evidence-based approach to this kind of evaluations, which must be tailored to specific groups of people (e.g., practitioners, patients) and contexts of use (Solomon et al. 2016).

Figure 2. The ranking of data item in radiology; items are enumerated according to their fine-grained ranking within the macro-category of priority level. Levels are established according to statistical testing.

DATA ITEM IN THE MAIN PAGE OF THE RADIOGRAPHER	LEVEL OF PRIORITY
Patient ID, Exam Date, Birth Date, Radiographist Name.	HIGHER
Level of urgency, Name of the patien, Time of execution, Patient Gender, Responsible Radiologist, Access number, Alert, Body section, DICOM ID, Patient fiscal code, Modality Type, Referring facility.	PROBABLY HIGHER
Ordering Radiologist, Exam description, Phone number, Room ID, patient's current address, patient's municipality, how the patient arrived, cost center.	PROBABLY LOWER
Patient nationality, Patient height.	LOWER

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

EXAMPLES OF USER STUDIES IN HEALTHCARE HDI

In this section we will report about three small user studies we accomplished under the tenets of HDI for an evidence-based design of data visualization tools for the healthcare sector. This is also aimed at showing that an evidence-based approach can be adopted also for small software development projects and without the need of relevant budgets. In particular, the first study regards primary data in radiology work, in terms of the identification of what data the practitioners should be required to record for their primary aim, and what data they deem less useful to this aim. The second study regards the value of tertiary data about hospital quality and availability, as this is evaluated by citizens that would need to gain information about the hospitals of their catchment area to choose where to go according to their needs and preferences. We will just outline the first two studies and give some more details of the third one, which regards the acceptance of an innovative system of data visualization for aggregate patient data by a sample of general practitioners (family doctors).

The Radiology User Study

In this study, we asked 15 expert technicians from different radiology facilities of Northern Italy to indicate the usefulness for their work of 31 information items on an ordinal scale, from 1 (negligible) to 6 (essential). The items regarded either the assessment of patients, their management for the execution of digital imaging examinations, or the preliminary evaluation of radiographic media, and they were collected by considering the Italian guidelines for quality assurance in teleradiology (Working Group for Quality Assurance in Diagnostic and Interventional Radiology, 2010).

The aim of the study was to rank those items from the most useful to the least, so that the interface of their information system could be optimized. Here, the concept of optimization should be interpreted as follows: either in terms of having the radiological information system display and require only those data that the radiographers considered certainly useful for their job; or, at least, in terms of a system that would not urge the users to fill in the low priority fields as mandatory data that block the application unless radiographers insert them. The online questionnaire platform displayed the 31 items in different order to minimize order bias and the task was carried out as part of an assignment of a first Level Master degree class on IT management for radiographers.

We then applied an original ranking method, discussed in (Cabitza et al. 2015a) by which we classified each item in one of four levels of priorities (see Figure 2). An informal interpretation of the result could consider the first group (see the green block in Figure 2), i.e., Patient ID, patient birth date, exam date, operator name, radiographer name (if these latter are not the same) as those items that cannot be missing in the radiographer screen, and that indeed need to be emphasized at the interface level (e.g. by being rendered at a proper font size) and double checked by the operators involved to avoid potential mistakes and adverse events. The second group of items (including the name of the patient, the ordering radiologist, the access number) are data of lower priority but still very important, that radiographers should consider mandatory to fill in. Conversely, low priority items like patient nationality and patient height (see the red block in Figure 2) could be considered those that should not be imposed to the data work of radiographers, nor displayed irrespective of the specific examination, as these unnecessary items could contribute to clutter the user interface and hide more relevant information.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In our future work, we will perform the same user study in different groups of users of the radiological information system, like the medical radiologists and the administrative clerks, so as to make the interfaces adaptive according to the role logged in, and have the systems display awareness promoting information (Cabitza & Simone, 2012) about the different perception of utility (if any) by different worker categories. For instance, let us assume that the administrative clerks in a given hospital would consider the indication of the cost center and the billing address of the patient essential data items for their paperwork; in this case, radiographers, who conversely attach a low priority to these items (see Figure 2), could be notified that these data are important for someone else in the same facility (e.g., by means of some iconic indication or by a warning message at the end of the work shift) and be made aware (see above) of the data-related needs of other professionals that are connected by the same patient case.

The Open Data Study

In this second study, we collected 330 complete questionnaires from a population of citizens living in an urban area and aged 18 or older. To this aim, we invited personal friends and acquaintances, all of the colleagues of our university department, and the students of two master classes from the same university. Sample representativeness was achieved by weighting the response set by both age and gender according to the latest national census. The respondents had to evaluate the utility of 11 information items in different scenarios of urgency and serious health conditions. The information items had been selected from those that can be conveniently extracted from some of the most important big open healthcare data sets available in the USA, in Canada and in Italy (taken, respectively, from healthdata.gov, www.cihi.ca and dovesalute.gov.it).

The sample exhibited (statistically) significant appreciation for three information items above all the others: the hospital ranking by number of admissions for the pathology of interest; the ranking by average wait time for the pathology of interest; and the ranking of facilities by their reaching time from one's place. The rankings evaluated the least useful were the rankings produced according to either the perceived level of care quality by all the hospitalized patients, or by the percentage of patients that had to be hospitalized again within six months since the first discharge. While the latter information is quite technical, perhaps too much to be appreciated by a sample of lay people, the former indication came unexpected and quite puzzling.

All of the collected indications are useful to design healthcare portals that show their users a minimal set of items with the highest potential to satisfy the users' needs, on default of other information, leaving the other items accessible on demand. Interestingly, we detected significant differences in the perception of information usefulness between different user profiles. For instance, male respondents attached more importance than women to the ranking based on georeferenced information. Not surprisingly, expert users considered the ranking by number of admissions for the pathology of interest more important than people self-declaring non-expert. We also probed the sample preferences for different ways to display the same information, either by simple and essential tables or by richer and more appealing infographics. Also in this case we detected a statistically significant difference between females and males, as the former ones found infographics to be clearer and more intuitive than men; and between young respondents and the elderly as the former ones, quit unexpectedly, found the tabular form more clear than the latter ones. These differences suggest that the usability of a healthcare portal can be maximized even considering basic profiling information and user satisfaction can be improved by asking the user only a few data, like gender and age.

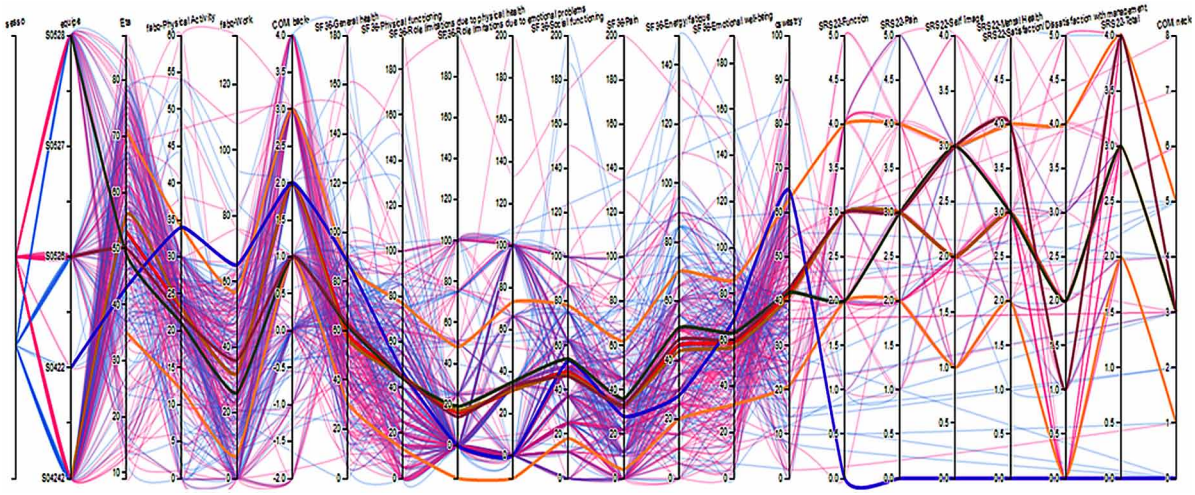
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The General Practice Study

We will focus now on the third user study that in some way integrates the research question of the former two studies regarding perceived utility with one of the central issues for the HDI field of whether interactivity improves the perceived value of data. Thus, the first research question is: What health open data are perceived as *more useful* in general practice (if any)? The second one is: Would visualizing these open data by means of *interactive* (that is browsable, zoomable, etc.) maps make these data to be perceived as *more useful and usable*? To further investigate this latter question we addressed it also in case of innovative visualizations, like in the case of diagrams displaying *parallel coordinates* (Inselberg, 2009) (see Figure 3). Thus, in particular, we wondered whether this kind of tools are effective means to represent patient data at various levels of description and to help GPs understand population trends, discover (frequency-based) normal and abnormal conditions, and detect correlations between multiple conditions.

To these above aims, we conceived a short questionnaire to be administered as a Computer-Aided Web Self-Interview (CAWI). This questionnaire consisted of four sections, which were to be rendered each in a different Web page: in the first page, the platform presented a list of 13 datasets that the authors had previously selected among the health open datasets made available on the DatiOpen.it platform (http://www.datiopen.it/en), which is one of the biggest Italian open data portals collecting more than 2,000 thousand datasets. In particular the respondents were asked to choose at least one dataset (and at most three) that they could want to consult in an online and customized "visual dashboard" supporting their own professional practice. In the second page we further selected 6 datasets for their suitability to be mapped into a georeferenced map like the one depicted in Figure 4.a, and asked the doctors to evaluate the degree of *utility for their profession* of each dataset on a six-value ordinal scale, from 1 (not useful at all) to 6 (very useful).

Figure 3. The parallel coordinate diagram shown to the user panel; the thick red and orange lines indicate the average values and standard deviations respectively. Blue lines represent single male patients. Pink lines female patients.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Then, we asked the doctors to imagine to be able to interact with the map, so that they could browse the map by dragging, and zooming in and out at the desired level of visualization (e.g., province, city, neighborhood if possible) and get a *heat map* like the one depicted in Figure 4.b. We then asked whether visualizing the above datasets on such a browsable map would make the data more useful or not (on a 5-value scale ranging from 1 "the data would be much *less* useful" to 5 "the data would be much *more* useful", passing through a middle value 3 "it wouldn't make almost any difference"). Finally in the third section, the questionnaire displayed a parallel coordinates diagram like the one depicted in Figure 3. We asked to indicate on a six-value ordinal scale the degree of informativity of such a way to represent patient data (from 1 not informative at all, to 6, very informative).

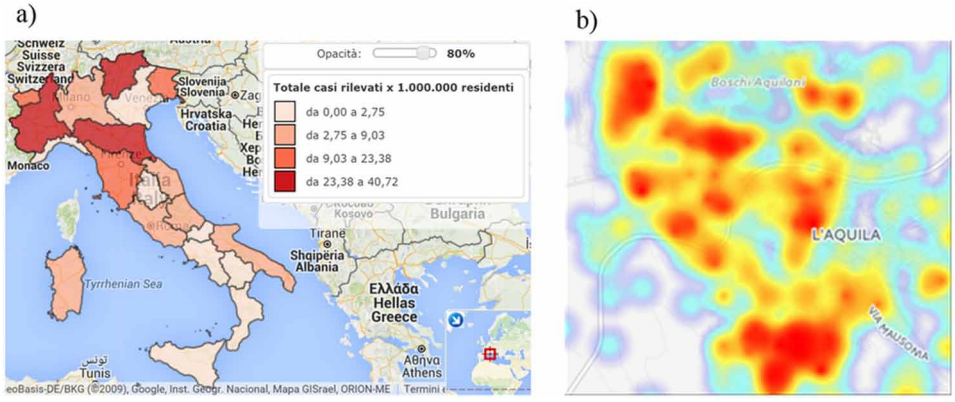
The questionnaire was left open for two weeks after the invitation had been sent to potential respondents. In this time lapse no reminder was sent. When we closed the survey, 29 GPs had accessed the questionnaire; 23 of them had completely filled in every item of the questionnaire in an average time of 5.9 minutes (SD=2.6 minutes). In this respondent sample, 5 out of 10 have been family doctors for more than 30 years, 9 out of 10 for more than 20 years; no respondent had less than 10 years of work experience so that we can consider the sample composed of experts in the Family Medicine domain.

The majority of the respondents (65%) claimed to be competent in Information Technology (IT) but neither experts nor enthusiasts. These latter ones were, respectively, approximately one third and one fifth of the sample. We detected a statistically significant and moderately negative correlation between age and IT expertise (-.49, p=.018), as it could be expected.

The respondents were asked to choose the three most useful datasets to be visualized on a personal dashboard for their daily practice by selecting them out of a list of 13 data sets. These had been selected for their popularity in terms of number of downloads from a comprehensive collection of Italian health open datasets . The sets chosen more frequently by the respondents were:

1. Regional distribution of the percentage of elderly people involved in the program of Integrated Home Care (in Italian: Assistenza Domiciliare Integrata).
2. Number (and details) of the Rehabilitation facilities for the elderly (in Italian: Istituti di Riabilitazione Extraospedaliera per Anziani).
3. Inpatient care average length of stay.

Figure 4. The static map (on the left, a) shown to the general practitioners; the heat map (on the right, b) resulting from interacting with the map depicted on the left



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

These datasets were selected by approximately one third of the sample (38%, 33% and 30% respectively). The other datasets collected much fewer preferences, with the "number of cases affected by pneumococcal disease by year" that was not chosen by any respondent.

In regard to the utility of data sets to be displayed on a georeferenced map, we got statistical significance for two sets only, which were considered useful for general practice, namely:

1. Number of registers on exposure to carcinogens (.5 vs. .95, p=.000).
2. Hospitalization rates, by regime, patient genre and region (.26 vs. .74, p=.035).

For the other data sets we did not get statistical significance: however, the "number of medical prescriptions delivered" can be considered the least appreciated information set (.64 vs. .36).

We also asked whether putting all of the above datasets into an interactive map, instead of a static one, would change the perceived utility of the datasets themselves. The majority claimed mapping data would make a difference, and that in so doing the datasets would be more useful (.83 vs. .17, p=.003, one third of the GPs even claiming that the data "would be much more useful"). The ranking of relative utility of mapping open data sets is reported in Table 1.

As a visualization aid to the comprehension of patient data and their correlation on multiple dimensions, parallel coordinates, like those depicted in Figure 3 and 5 were found to be highly informative: the responses were positive for the majority (.64 vs. .46, p=.286). Furthermore, almost every GP involved in the user study claimed that the interactive version (see Figure 5) of the diagram was much more informative than the static view (.96 vs. .04, p=.000, 60% "much more informative").

The characteristics of the respondent sample, in terms of work experience and attitude towards Health IT suggest that the respondents can be considered representatives of a population of real domain experts and that the potential (positive) bias introduced by the IT enthusiasts should be low. Besides the results reported above, the questionnaire also allowed the respondents leave free text comments in order to suggest datasets and applications to be considered in further studies, give advice on potential improvements to the visual tools presented, and more generally share remarks with the research team.

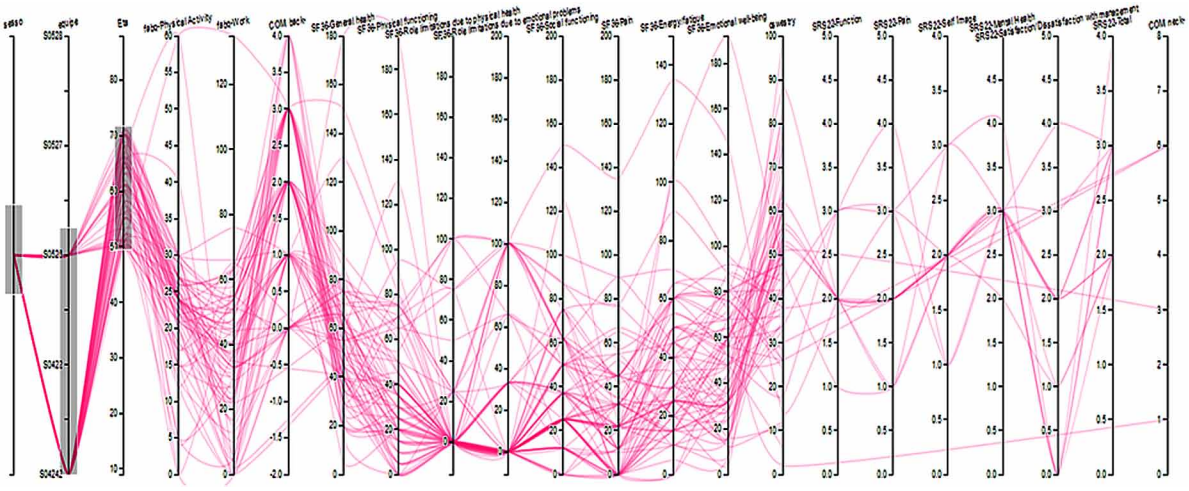
In particular, we report two comments that suggested relevant improvements to the Parallel Coordinate (PC) dashboard. One GP suggested to implement PCs indicating explicitly central tendency parameters on their axes (wherever applicable), like means and medians, standard deviations and interquartile ranges,

Table 1. Open datasets to be visualized in maps ranked by perceived utility

Dataset	Priority level$^{(sig.)}$	Perceived utility$^{(sig.)}$	Median
Carcinogen exposure registers	Higher$^{(*)}$	Positive$^{(***)}$	5
National Drug Consumption	Uncertain$^{(NS)}$	Positive$^{(NS)}$	4
Home care patients	Uncertain$^{(NS)}$	Positive$^{(NS)}$	4
Hospitalization rate	Uncertain$^{(NS)}$	Positive$^{(*)}$	4
Incidence of invasive diseases	Uncertain$^{(NS)}$	Positive$^{(NS)}$	4
Incidence of occupational diseases	Uncertain$^{(NS)}$	Positive$^{(NS)}$	4
Number of prescriptions	Uncertain$^{(NS)}$	Negative$^{(NS)}$	3

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The parallel coordinate diagram shown in Figure 3 after having interacted with it; in particular three specific ranges of data have been specified to filter the patient data to be highlighted: more precisely, female patients treated by two specific medical teams (equipe) and between 70 and 50 years of age



both at the level of the single GP patient set, and (if available) at the regional and national level. This doctor noticed how the bundle of lines, in virtue of their variable thickness at the intersection with the different coordinates, would nicely and visually represent the frequency distribution of the values within the subject population. Another GP suggested to allow the user to switch from a cross-sectional view (i.e., the default one) where each line represents the current conditions of a single patient, to a longitudinal view. In such a view, the current conditions would be represented as a highlighted line, while the bundle of the other lines would represent the different conditions that the same patient exhibited in the past, as these are recorded on the medical record of the GP or in the patient's health record.

We are aware of the limitations of this study. For instance, the fact that the involved doctors had to imagine the novel interaction, by looking at static pictures of how the visualization tools would change according to their actions. However, giving doctors real dashboards with which to interact and ask about their perceptions would have been problematic for the impossibility to give them a uniform training and to check remotely their interaction with the tool to evaluate. Using an interactive prototype would have implied a training phase of all the doctors involved in the experiments, and a physical presence during the usage of the interactive prototype to observe and record the experiment for further examination. A uniform training phase could have being achieved by preparing a video animation to show how to use the interactive prototype, but we could not do it due to two main reasons: (i) the lack of human resources that should also be present during the experimental sessions, and hence the impossibility to control the experiment adequately; (ii) the dread that the respondents, acting remotely without our presence, could assess the video animation and not the interactive prototype. However, we believe that studies like the present one, where interactivity is simulated in terms of its effects, could be useful anyway in a preliminary phase of development, where to understand if endowing visual tools with some kind of interactivity would be worth the effort.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

In light of the evidence collected in this study, we can conclude that designing for a better interactivity of open data platforms for general practice is worth further efforts. From the HDI perspective, this conclusion was desirable; although it might also seem easily predictable, nevertheless that conclusion could not be taken for granted in a committed *evidence-based approach* to Health IT design (Rigby et al. 2013, Longhurst et al. 2013, Cabitza et al. 2016). For this reason, our studies are just a small but necessary contribution in the research agenda outlined in the Introduction and that we advocate as future research directions of an increasing number of researchers in the health informatics, data science and human computer interaction fields. These future researches should be aimed at gaining feedback and design-oriented indications from both caregivers and patients to improve their experience in interacting with the next data visualization tools to come.

CONCLUSION

In this chapter we have proposed a tripartite perspective to health data by which to distinguish between primary data, i.e., the content of health records that clinicians and nurse produce and consume in their decisions and treatments; secondary data, i.e., the content of structured data bases, data warehouses and the reports that automated procedures produce and display on interactive data visualization tools (or dashboards) to support the tasks of clerical and managerial roles (including policy makers); and tertiary data, i.e., particular content that has been selected and optimized in terms of clarity, understandability and appeal, in one word usability, to be offered online for the general public and hence for unanticipated needs and aims. The distinction has been proposed to separate concerns and differentiate agendas. The one-size-fits-all approach to data quality improvement simply does not work (Pipino et al. 2002), nor solve the still present problems of low use of and low satisfaction for the Electronic Health Record (Cabitza et al. 2015a). Moreover, such a simple separation of concerns also allows to emphasize the concept of "use and reuse" of health data (Berg and Goorman, 1999), which is important to keep in mind when setting quality requirements and imposing ideal data model to health practitioners and their patients.

Human Data Interaction (HDI), as a research approach, can provide alternative methods to address these challenges, which are becoming more and more urgent especially in our current age of increasing opportunities and threats regarding the "datafication" of medicine and the consequent production of health big data (Murdoch & Detsky 2013). In particular, HDI approaches would focus on evaluating how ICT supports data (producing) work by relieving doctors and nurses from form-filling, coding and other administrative tasks; and on evaluating the usability and the user experience related to interacting with tertiary (health) data, to understand the actual value these convey in decision making and information-intensive tasks. Put it in another way, through HDI a design approach more sensible to relieving doctors from administrative tasks should become the primary concern behind ICT systems for supporting healthcare practices. It goes without saying that there is not an ultimate solution to achieve this goal, and HDI is no exception. Nevertheless, HDI may promote an attitude and awareness to cognitive ergonomics aspects, to usability studies applied to healthcare information systems, and so on. The HDI perspective is aimed to let emerge that doctors should not work with data as information *per se* (e.g., by forcing them to fill a form of a database table full of tedious completeness constraints aimed to achieve an intrinsic data quality), rather they should use their time to work and reflect *upon* data, in order to enhance their

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

daily work, practices and decisions. As a final consideration, HDI may be considered as an approach through which nothing is advisable *a priori*, unless it is packed-up with an evidence-based design and empirical evaluation. HDI cannot be directly connected to better outcomes, but to trace back all of the ICT recommendations and solutions to the evaluation of their impact in daily practices in terms of outcome efficiency and effectiveness, rather than in terms of intrinsic characteristics of a data system *per se*.

REFERENCES

Abdelhak, M., Grostick, S., & Hanken, M. A. (Eds.). (2012). *Health information* (4th ed.). Elsevier.

Abiteboul, S., & ... . (2000). *Data on the Web: From Relations to Semistructured Data and XML*. San Francisco, CA: Morgan Kaufmann Publishers.

Ammenwerth, E., & Rigby, M. (Eds.). (2016). *Evidence-Based Health Informatics: Promoting Safety and Efficiency Through Scientific Methods and Ethical Policy* (Vol. 222). IOS Press.

Authors, V. (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, *268*(17), 2420. doi:10.1001/jama.1992.03490170092032 PMID:1404801

Berg, M. (1999). Accumulating and coordinating: Occasions for information technologies in medical work. *Computer Supported Cooperative Work*, *8*(4), 373–401. doi:10.1023/A:1008757115404

Berg, M., & Goorman, E. (1999). The Contextual Nature of Medical Information. *IJMI*, *56*, 51–60. PMID:10659934

Berg, M., & Toussaint, P. (2002). The mantra of modeling and the forgotten powers of paper: A sociotechnical view on the development of process-oriented ICT in health care. *JMI*, *69*(2-3), 223–234. PMID:12810126

Bossen, C. (2016). Data-work in Healthcare: The New Work Ecologies of Healthcare Infrastructures. *ACM International Conference on Computer-Supported Cooperative Work and Social Computing, CSCW 2016*. 10.1145/2818052.2855505

Cabitza, F., & Batini, C. (2016). Information Quality in Healthcare. In Data and Information Quality, (pp. 421–438). Springer. doi:10.1007/978-3-319-24106-7_13

Cabitza, F., Del Zotti, F., & Misericordia, P. (2014). Electronic Records for General Practice – Where we Are, Where we should Head to Improve Them. In *Healthinf'14: Proceedings of the International Conference on Health Informatics*, (pp. 535-542). INSTICC.

Cabitza, F., Locoro, A., & Batini, C. (2015b). A User Study to Assess the Situated Social Value of Open Data in Healthcare. *Procedia Computer Science*, *64*, 306–313. doi:10.1016/j.procs.2015.08.494

Cabitza, F., Locoro, A., Fogli, D., & Giacomin, M. (2016). Valuable Visualization of Healthcare Information: From the Quantified Self Data to Conversations. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 376-380). ACM. 10.1145/2909132.2927474

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cabitza, F., & Simone, C. (2012). Affording Mechanisms: An Integrated View of Coordination and Knowledge Management. *CSCW*, *21*(2), 227–260.

Cabitza, F., Simone, C., & De Michelis, G. (2015a). User-driven prioritization of features for a prospective InterPersonal Health Record: Perceptions from the Italian context. *Computers in Biology and Medicine*, *59*, 202–210. doi:10.1016/j.compbiomed.2014.03.009 PMID:24768267

Cafaro, F. (2012). Using embodied allegories to design gesture suites for human-data interaction. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp 2012*, (pp. 560–563). ACM. 10.1145/2370216.2370309

Crabtree, A., & Mortier, R. (2015). Human data interaction: Historical lessons from social studies and CSCW. In *ECSCW 2015*. Springer. doi:10.1007/978-3-319-20499-4_1

Elmqvist, N. (2011). Embodied human-data interaction. ACM CHI 2011 Workshop "Embodied Interaction: Theory and Practice in HCI", 104–107.

Fitzpatrick, G. (2004). Integrated care and the working record. *Health Informatics Journal*, *10*(4), 291–302. doi:10.1177/1460458204048507

Fitzpatrick, G., & Ellingsen, G. (2013). A review of 25 years of CSCW research in healthcare: Contributions, challenges and future agendas. *CSCW*, *22*(4-6), 609–665.

Gitelman, L. (2013). *Raw data is an oxymoron*. MIT Press.

Gregory, J., Mattison, J. E., & Linde, C. (1995). Naming notes: Transitions from free text to structured entry. *MIM*, *34*(1-2), 57–67. PMID:9082139

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition*, *84*(3), 343–352. doi:10.1016/S0010-0277(02)00050-1 PMID:12044739

Hornung, H., Pereira, R., Baranauskas, M. C. C., & Liu, K. (2015). Challenges for Human-Data Interaction–A Semiotic Perspective. In *International Conference on Human-Computer Interaction* (pp. 37-48). Springer International Publishing.

Inselberg, A. (2009). Parallel coordinates. In L. Liu & T. Ozsu (Eds.), *Encyclopedia of Database Systems* (pp. 2018–2024). Springer.

Kennedy, J. B., Mitchell, K. J., & Barclay, P. J. (1996). A framework for information visualisation. *SIGMOD Record*, *25*(4), 30–34. doi:10.1145/245882.245895

Liaw, S. T., Rahimi, A., Ray, P., Taggart, J., Dennis, S., de Lusignan, S., ... Talaei-Khoei, A. (2013). Towards an ontology for data quality in integrated chronic disease management: A realist review of the literature. *IJMI*, *82*(1), 10–24. PMID:23122633

Lieberman, H., Paternò, F., Klann, M., & Wulf, V. (2006). End-User Development: An Emerging Paradigm. EUD, 9, 1–8.

Locoro, A. (2016). A Map is worth a Thousand Data: Requirements in Tertiary Human-Data Interaction to Foster Participation. *CEUR Workshop Procs of CoPDA 2015*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Locoro, A., Grignani, D., & Mascardi, V. (2011). MANENT: An infrastructure for integrating, structuring and searching digital libraries. *Studies in Computational Intelligence, Springer*, *375*, 315–341.

Longhurst, C. A., Palma, J. P., Grisim, L. M., Widen, E., Chan, M., & Sharek, P. J. (2013). Using an Evidence-Based Approach to EMR Implementation to Optimize Outcomes and Avoid Unintended Consequences. *Journal of Healthcare Information Management*, *27*(3), 79. PMID:24771994

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Journal of the American Medical Association*, *309*(13), 1351–1352. doi:10.1001/jama.2013.393 PMID:23549579

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors: Hum. Factors*, *52*(3), 381–410. doi:10.1177/0018720810376055 PMID:21077562

Pipino, L., & Lee, Y. W. (2011). Medical Errors and Information Quality: A Review and Research Agenda. AMCIS.

Pipino, L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, *45*(4), 211–218. doi:10.1145/505248.506010

Rigby, & Ammenwerth, Beuscart-Zephir, Brender, Hypponen, Melia, … de Keizer. (2013). Evidence Based Health Informatics: 10 years of efforts to promote the principle. *Yearbook of Medical Informatics*, *8*(1), 34–46. PMID:23974546

Roberts, A., Gaizauskas, R., Hepple, M., & Guo, Y. (2008). Mining clinical relationships from patient narratives. *BMC Bioinformatics*, *9*(11), S3. doi:10.1186/1471-2105-9-S11-S3 PMID:19025689

Sauleau, E. A., Paumier, J.-P., & Buemi, A. (2005). Medical record linkage in health information systems by approximate string matching and clustering. *BMC Medical Informatics and Decision Making*, *5*(1), 32. doi:10.1186/1472-6947-5-32 PMID:16219102

Shankaranarayan, G., Ziad, M., & Wang, R. Y. (2003). Managing data quality in dynamic decision environments: An information product approach. *Journal of Database Management*, *14*(4), 14–32. doi:10.4018/jdm.2003100102

Shortliffe, E. H., Perreault, L. E., Wiederhold, G., & Fagan, L. M. (Eds.). (1990). *Medical Informatics: Computer Applications in Health Care and Biomedicine*. Addison-Wesley.

Smith, P. C., Araya-Guerra, R., Bublitz, C., Parnes, B., Dickinson, L. M., Van Vorst, R., & Pace, W. D. (2005). Missing clinical information during primary care visits. *Journal of the American Medical Association*, *293*(5), 565–571. doi:10.1001/jama.293.5.565 PMID:15687311

Solomon, J., Scherer, A. M., Exe, N. L., Witteman, H. O., Fagerlin, A., & Zikmund-Fisher, B. J. (2016). Is This Good or Bad? Redesigning Visual Displays of Medical Test Results in Patient Portals to Provide Context and Meaning. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2314-2320). ACM. 10.1145/2851581.2892523

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM*, *40*(5), 103–110. doi:10.1145/253769.253804

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Swinglehurst, D., Greenhalgh, T., & Roberts, C. (2012). Computer templates in chronic disease management: Ethnographic case study in general practice. *BMJ Open*, *2*(6), e001754. doi:10.1136/bmjopen-2012-001754 PMID:23192245

Timmermans, S., & Berg, M. (2003). *The gold standard: the challenge of evidence-based medicine and standardization in health care*. Temple University Press.

Wang, R. Y. (1998). A product perspective on total data quality management. *Communications of the ACM*, *41*(2), 58–65. doi:10.1145/269012.269022

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *JMIS*, 5-33.

ADDITIONAL READING

Coiera, E. (2004). Four rules for the reinvention of health care. *BMJ: British Medical Journal*, *328*(7449), 1197–1199. doi:10.1136/bmj.328.7449.1197 PMID:15142933

Goetz, T. (2010). *The decision tree: taking control of your health in the new era of personalized medicine*. Rodale.

Greenhalgh, T., Howick, J., & Maskrey, N. (2014). Evidence based medicine: A movement in crisis? *BMJ (Clinical Research Ed.)*, *348*, g3725. PMID:24927763

Munzner, T. (2009). A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 921–928. doi:10.1109/TVCG.2009.111 PMID:19834155

Shneiderman, B., Plaisant, C., & Hesse, B. W. (2013). Improving healthcare with interactive visualization. *Computer*, *46*(5), 58–66. doi:10.1109/MC.2013.38

KEY TERMS AND DEFINITIONS

Data Visualization: A visual representation of a (subset of) dataset. Also, the scholarly discipline addressing how to design, implement and deploy this kind of representations.

Data Work: A class of working activities that regard the production, arrangement, management and interpretaton of data to perform other work tasks adequately.

Evidence-Based Design: A design approach that integrates the empirical and user-centered validation of alternative solutions to choose which ones to implement and how to, in order to increase the odds that the overall final system will be found to be usable (efficient, effective, satisfactory) by the users and be appropriated in their practices.

Human-Data Interaction: a discipline that addresses the phenomenon of how users make value out of datasets by interacting (i.e., querying, filtering, zooming, etc.) with an either textual or visual representation of those data.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Open Data: Those non-personal data that, made available by some institution, organization or association, anyone can freely access, use, modify, and share for any purpose, subject, at most, to requirements that preserve attribution and openness.

ENDNOTES

1 We concur with Hornung et al. (2015) that data are "artifact-mediated representations of phenomena that need to be given meaning by people and that serve some purpose".

2 To this respect, datafication is a process of both dis-closure of the representable reality, as well as the closure of what is (willingly or unaware) left tacit behind. Datification constructs facts, but it can also entail the (irreversible) destruction of context, through the emergence of text out of human experience.

3 Obviously we are aware that the value of something, and hence also data, can be appraised differently than by adopting a consumer-oriented perspective, which regards what in classical political economy is denoted as "use value". Value can be a function of the resources involved or consumed to get the data (mainly by primary users), to verify and cleanse them, and make them fit to interoperable formats typically in secondary processes); or, simply, a function of how much someone is going to pay to obtain them (cf. the idea of economic value, that is so common in this age of splendour of online advertising). These different perspectives reflect different conceptions of the idea of value, as it has been formulated in centuries of economic and theoretical speculation (see (Dobb, 1975) for a classic review of those conceptions).

4 Other interesting cases regard the failure to detect relations between either right data or inconsistent ones; the failure to get the right implication from right data or, the other way round, success in detecting right relations and inferring right implications from wrong data. It is important to distinguish between all of these cases in understanding the root causes of information failures and understand how to minimize the odds of their occurrence.

This research was previously published in Smart Technology Applications in Business Environments edited by Tomayess Issa, Pedro Isaías, Touma B. Issa, Piet Kommers, and Theodora Issa; pages 184-203; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 59
An Optimized Graph-Based Metagenomic Gene Classification Approach: Metagenomic Gene Analysis

Md Sarwar Kamal
East West University, Bangladesh

Mohammad Ibrahim Khan
Chittagong University of Engineering and Technology, Bangladesh

Kaushik Dev
Chittagong University of Engineering and Technology, Bangladesh

Linkon Chowdhury
Chittagong University of Engineering and Technology, Bangladesh

Nilanjan Dey
Techno India College of Technology, India

ABSTRACT

Biological interaction mainly depends on the interactions of various genes and genomes. To identify actual meaning of interactions we have to find out the facts and reasons for these interactions. Gene analysis allows to verify such environment. Gene annotation means to identify the exon regions in metagenomic samples. The de Bruijn graph plays significant role in gene prediction and next generation sequencing (NGS). Apart from that, Eular Path of de Bruijn graph introduced generalized gene annotation for translational and splicing signals, exon introns separation and coding regions. set of graph reduction rules have used to build a de Bruijn graph. Accurate solution for large scale sequencing, trims space complexity and generates optimal gene annotation have tested.

DOI: 10.4018/978-1-7998-1204-3.ch059

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

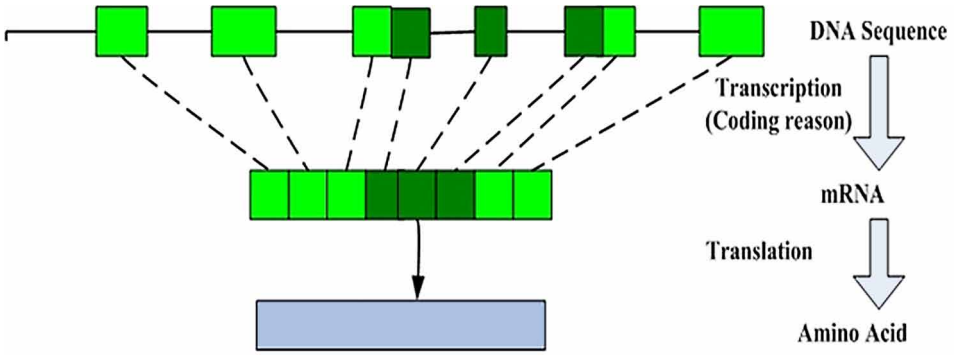
INTRODUCTION

The approach that specifies or categorized the genetic substance from environmental collections is called metagenomics. This process is important due to its uniqueness on data clustering facilities. I.e. the system permits to organize huge molecular data into a small area. By imposing metagenomics concept, various vital and unique molecules with pivotal activities and applications have been determined. The populations and demand of foods and daily necessary goods are increasing day by day. To fulfill the demand of daily commodities as well as medical support, metagenomics will play very important role. Lot of companies are willing and investing in this sector to clear the unknown genetic information. So metagenomics gene analyses will dominant in enzymes processes, medicines, agriculture, ecology and new generic products. Metagenomics ensure to supply new molecules and novel enzymes with dynamic activities and provide factors from set of enzymes from the cultivable microorganisms. Apart from that, the uses of metagenomics are used to uncovering novel biocatalysts from nature, bioremediation and xenobiotic metabolism.

DNA sequence consists of four nucleotides A, G, C and T. These sequencing efforts generate a large amount of raw data as the DNA sequence of a eukaryote is often longer than a hundred million base pairs. The genome of human has 3.2 billion (appx) base pairs. The annotation of these sequences of biological data is needed as a computation tool for gene finding. Locating the genes is not only helpful but also a prerequisite for further analysis, such as the characterization of the function of the gene product, determining the phylogeny of different species or understanding gene regulation. But the problem of finding genes in a genomic DNA sequence is difficult and high space complexity.

Metagenomics, the investigation of crude bacterium under the study of genomic substances gathered precisely from natural data sets, has been applied to analyze the shape, activities, and mechanical interactions of small communities. Various analysis permit us to know about different microbes in multiple areas (Ley et al., 2008; Rusch et al., 2007; Wu et al., 2011), their composite architectures of the communities they built (Huttenhower et al., 2012), and the pivotal encounter they maintain on the associated environment (Tyson et al., 2004). It important to analyze the various bacterium interactions

Figure 1. The central dogma is a framework for understanding the flow of genetic information. It states that DNA makes RNA, and RNA makes protein (amino acid.)In transcription process DNA convert into RNA specially mRNA(messenger RNA) and in translation RNA convert into protein(amino acid).



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

however, human metagenomics interactions are very essential to utilize the meaningful result due to its impact on mankind need as drugs, foods and diseases identifications (Qin et al., 2012; Behre C.J, et al., 2013; Greenblum et al., 2011).

In the age of information superhighway, numerous metagenomic analyses are based on gene-centric processes and as well as new aim to cluster the complete genomic shape of the association by virtue of shotgun metagenomics. Mathematical analysis based on shotgun sequencing impact helped to mapped database to a orthologous gene groups such as KEGG (Kanehisa et al., 2012), COG (Tatusov et al., 2003), EggNOG (Powell et al., 2012), M5NR (wike et al., 2012) to investigate similarities to genes or proteins with real and annotated activities. It is important to know that the biological activities among similar genes are work together in contrast, with taxonomic identify resulted in the experimental analysis. BLAST is frequently used to mapped to obtain greater score values on sequence alignment and annotations (Dalevi et al., 2008). This annotation mechanism and the measurements of the data set alignment that mapped individual activity is a an important part in computational metagenomic research and it helps to identify the functional shifts related with disease, as it is the basis of general genomic of analyses.

An important hypothesis afterwards, this annotation process and the key to properly saving the absolute structures of the metagenome alignment where small genome sequencing input holds abundant information to precise map to the exact activity. By considering different annotation rules and assumptions that each input will be directed to either its gene of or to an orthologous gene from an efficient distant genome annotated with the same function. These rules for future assumption that input initiating from intergenic regions or from orthology collection with unknown activity will not be properly mapped to any of the active annotated genes in our database. These considerations are vital for both estimation as well as identifying the presence or absence of gene groups. Now-a-days, but, no details and systematic studies has not checked this analysis on a broad scale, nor have rules for this annotation system been regulated. The shortage of such systematic analysis produce the results sketched for different metagenomics studies are very difficult to quantify and compile and challenges any steps to compare research outcomes.

Microarray data set is a collection of chemically coated glass, nylon membrane or silicon. A microarray is a small chip (made of chemically coated glass, nylon membrane, or silicon. It comprises of millions of information for metagenomics genome or DNA sequence analysis. Metagenomics based research or microarray engineering organized an environment to measures automated process to check up the expressions and interactions of genes in all respects as drug design, disease investigations, protein-protein interactions and sequence analysis. But, very large data set in genes and biological network creates computational complexities in efficient data handling. To overcome the overcome the difficulties, classification and clustering techniques help optimize data processing. Subgroups help to organize the data set separately and consecutive ways. This clustering process is important for better data processing. Various clustering algorithms have been contributing the better data processing. A microarray based computational algorithm (Jiang et al., 2004) leveled data set into three distinct parts that helps to maintain group based data separation. But unsupervised data management produces some pitfall of this algorithm. Total number of groups and subgroups were unknown it becomes problematic for very large data processing. Microarray DNA sequence and gene analysis plays an important role towards the generations of drugs and diseases investigations. This system allows thousand of small DNA molecules in a small storage area. In current time, microarray data processing is very important in the field of Computational Biology. Consequently, microarray image processing enables proper information generation in gene annotations (Karim et al., 2011). Moreover, complete or global metagenomics processing opens verification techniques for gene expressions interactions. Formal and basic approaches of measuring gene study focused on a specific

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

factor that led a limited area of biological data processing as well as information generation. Due to limitations, lots of vital biological changes were lost or misjudged. In this regards global data processing are very much efficient to cover all the facts related to the overall data set (Fryer et al., 2002).

In addition, de bruijn graph runs much faster and requires less memory space than most of compared methods, and exhibits competitive CPU time. The performance of optimized de bruijn graph is even comparable with the other gene assembler in both sensitivity and accuracy. Reduced de bruijn graph also runs much faster for gene annotation.

BACKGROUND

De-Bruijn graph is a directed graph that crates edges among all existing nodes in such a way that $P = (p1, p2, p3, \ldots, an)$ to $Q = (q1, q2, q3, \ldots, bn)$ with Q being a left sided part of P, or $q1 = p2$, $q2 = q3$. The critical benefit of this graph is that it mathematical tools that helps to remove assembly faults and get optimal solutions. Transcription assembler iterative De Bruijn graph approach (T-IDBA) (Peng et al., 2011) is an modified approach of iterative de Bruijn graph that rearrange positive genome segments from large sets of irregular data set of genes. Velvet (Zerbino at el., 2008) is a complete tool that quickly manage de Bruijn graph to curtail false reads and cut the redundant input sequences in genome segments. Velvet is a strong and efficient and simple dBG-oriented assembler. ABySS (Simpson et al., 2009) is a parallel and cluster based a measurement which is used to employ storage space for large-scale data volumes using the de Bruijn based classification. This package support and indicate the K-mer graphs and strongly connected graph calculation for a certain amounts of nodes that requires large storage area. AllPaths (Butler et al., 2008) is another de Bruijn Graph based assembly tool which can manage large amounts of genome sequencing. It is expandable in both prototype and real data sets. Apart from that, this package holds an algorithm that is used to read corrected input genome sequences. An organized research paper (Sonia and Sarwar, 2015) on Next Generation Sequencing has narrated the graph based approaches for genome assembly. MIRA3 is an organized framework of overlap layout consensus (OLC) under greedy algorithms (Zang et al., 2000). It is a tedious assembler that can read and manage older errors by the support of OLC and greedy mechanisms. Mosaik (http://bioinformatics.bc.edu/marthlab/Mosaik) is a history oriented sequencing process that aligns parallel sequencing in an exact analysis using Smith–Waterman algorithm with the help of hash-based table for shorter sequence reads. Another OLC-based automated tool is Newbler (Margulies et al., 2005) which is massively re-arranged by the famous 454 Life Science in gene sequencing and analysis. The starting point of this tool is used to sequence the DNA or RNA without the support of pairwise alignments where the sequence lengths were near about 120 bp. This automated computational tool has been updated that can organize the pairwise alignments as well as distinct alignments. From 2005, OLC was introducing for two basic ways as, OLC could build unitigs from input data and second it constructs huge contigs from unitigs. The pivotal benefits of Newbler are uses of sequence reads coverage and reduced matching errors. Graphlet degree vectors (Sonia and Sohel, 2014) are tinny strongly connected level graph under a large graph network. It supports to verify the gene annotations as well as interactions.

The research area of finding the protein coding genes in a eukaryotic genomic sequence by computational methods has become competitive. Most of the estimated 30000-40000 human genes code for proteins. Cells make a protein molecule from a DNA sequence containing the gene (Figure 1). First, a copy of a contiguous region of DNA is larger than actually coding part. This process is called transcription.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Out of this sequence the introns are spliced out and the result is the concatenation of the exons, so called messenger RNA (mRNA). The number of exons varies in per gene. Some genes consist of only one exon and therefore do not contain introns. The human muscle protein constitutes the other extreme. It consists of 178 exons. In human, it is estimated that there are on average nine exons in per gene (Brown, 2002). In metagenomic analysis is defined as the classification of microbial genomes via the direct isolation of genomic sequences from the environment without prior cultivation (Handelsman, 2007).

The whole genome of an organism sequence is assembled and then genes are predicted. In metagenomic, a single genome cannot be assembled and therefore, it is a challenge to predict genes within these short sequences (Hoff et al., 2009). Accurate gene annotation for environmental samples is needed so that genes can be classified to their correct functions, and it paves the way for functional studies in metagenomic. In the chromosomes the DNA sequence is found in a double stranded sequence, where each nucleotide in one strand is opposed by a complementary nucleotide in the other strand. A(Adenine) is the complement of T(Thiamine) and vice versa. In the same way C(Cytosine) and G(Guanine) are complementary bases. As the sequence of nucleotides of one strand is fully determined by the sequence of nucleotides on the opposite strand, we only need to know the sequence of one of the strands. Usually two neighboring genes (regardless whether they are on the same strand or on opposite strands) are separated by an inter-genomic region, but exceptions to this organization are known. In the human mitochondrial genome the coding regions (sic) of the genes ATP8 and ATP6, which are on the same strand, overlap by 46 nucleotides (the reading frames in the overlapping region are different). An example of a gene lying within another gene is the human neurofibromatosis type I gene on chromosome 17, which has three short genes on the opposite strand within one of its introns. Each of these internal genes has introns itself.

LITERATURE OVERVIEW

Bridger (Chang et al., 2015) is a software package that calculates the genome arrangements with the help of de novo transcriptome assembler where Cufflinks based methods are employed to decrease drawback of ongoing de novo assemblers. Justification of this tool are made with real data sets of human, and mouse RNA. It reduces the errors of false positive selections and operates substantially faster and holds much less storage areas than most existing assemblers. Consequently, Bridger restricted the sensitivity and accuracy with Cufflinks. However, this process is not fit for handling overlapped sequences. VirAmp (Yinan et al., 2015) is an integrated assembler that comprised with the techniques of existing assemblers as well as web-based interactive interfaces. This assembler supports multiple pipeline environments for data set classifications. This package also enables representations of large number of data set into a single platform. The interactive interface is the key advantages of this tool. However, this tool is also not free from overlapped genome sequencing and it takes more time to handle interconnected genome sets. ClusDCA (Wang et al., 2015) is ontology based automated environment that relocate information among all data set that has unique activity to mitigate the overfitting difficulties for sparsely annotated functions. To depict interconnected data, ontology based method takes more time due to absence of mapping. A mapping based analysis overcomes the difficulties where a reads mapping algorithm utilize short reads onto a de Bruijn graph of assemblies (Yuzhen et al., 2015). The prime consideration is a hash table that organizes the relationship of k-mers (mer: from Greek meaning "part") to enable faster mapping of reads to the graph. This is a reference-centric technique to metatranscriptome assembly that utilizes graphs of metagenome assembly as the reference. The experiments however notice that, this process has

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

missed mapping of shorter fraction of multi-edge-spanning reads. While there is a sequencing defect in classification of k-mer, this algorithm is unable to detect the location of the error.

Accurate gene annotation for sample data is needed so that genes can be classified to their correct functions. Gene prediction programs can be categorized into two different groups. First, the ab initio programs that employ training set with known annotation for training the parameters of their models, are widely used in gene prediction (Mario et al., 2003). There are a large number of gene-finding programs of this kind, e.g. GENIE,(Martin et.al., 2000) GENSCAN, (Christopher, 1997) GENEID, (Gens et al., 2000) and GLIMMER (Arthur et al., 2007).Other gene prediction program known as similarity-based programs, use external information about the input sequence to predict genes. Some popular similarity based programs are GENEWISE,(Ewan et al., 2004) AGenDA (Leila, 2003) and the well known BLAST which searches the input. Open Reading Frame (ORF) is a conventional approach in gene prediction which began with a start codon and end with in-frame stop codon (Hoff et al., 2009). Generally, genes in prokaryotes are 1000-bp in length on average (Xu et al., 2006). Due to the short sequence length of metagenomic, they contain incomplete ORFs that lack start and/or stop codon, thus conventional ab-initio programs cannot be applied to metagenomic. Similarly, homology-based approaches for gene predictions rely on databases that only contain known, and thus a limited set of genes. Therefore, both of these categories do not work well for metagenomic fragments which are about 700 bp less than 400 bp when produced by Sanger and next generation sequencing, respectively (Mardis, 2008).Sequence similarity search has been proved useful in many cases, it is seen that only a fraction of newly discovered sequences have identifiable homolog in the current databases. Proteins might have already included representatives of most ancient conserved regions (ACRs, regions of protein sequences showing highly significant similarity across phyla) and that new sequences not similar to any database sequence are unlikely to contain ACRs (Green et al.,1993). The proportion of vertebrate genes with no detectable similarity in other phyla is estimated to be 50%.This is supported by a recent analysis of human chromosome 22 where only 50% of the proteins are found to be similar to previously known proteins. Therefore, recent tools have emerged to address these problems for metagenomic reads. Three programs are widely used for this purpose: Orphelia (Hoff, 2009), MetaGene (MG) (Noguchi et al., 2006), Meta Gene Annotator (MGA) (Noguchi et al., 2008), and Gene-Mark (Besemer et al., 1999). We demonstrate a memory efficient technique for metagenomic large scale data. We used bipartite graph and Euler path for finding gene from large DNA data set.

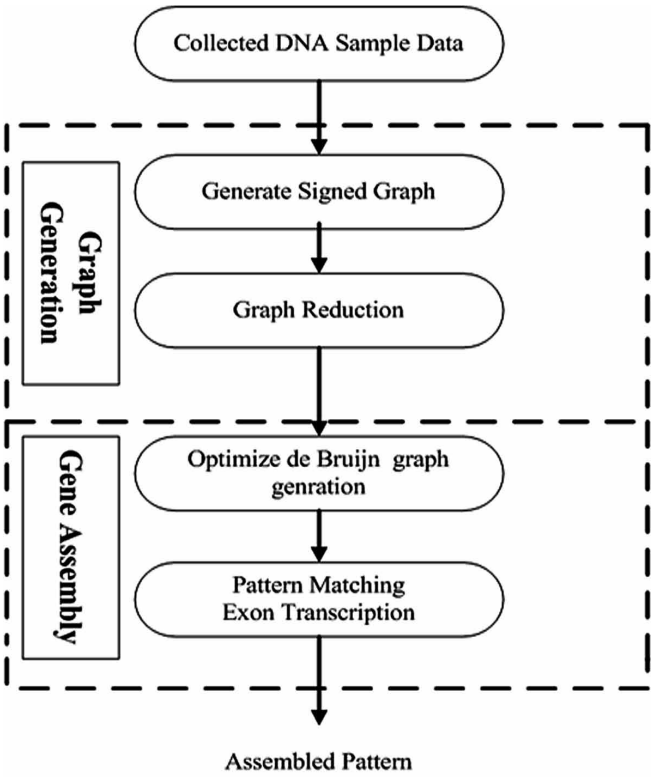
ARCHITECTURAL APPROACH

The structure of this work is built based on gene annotation (Figure 2).Our method works based on three phases: Data collection, Graph generation and Gene Annotation. The sample data are the data set of DNA nucleotides of human, plants or micro organ. Sample data are character of set of DNA nucleotide data stream. In graph generation phase divided into sub phases: generation of signed graph and graph reduction. In graph generation phase, we generate a undirected sign graph with multiple edges and loop. Graph reduction rules are used for graph rewriting. Basic reduction rule for graph rewriting technique is used for specification and generation of graph optimization.

After graph optimization we performed gene annotation. In next phase, gene annotation process is applied with euler path by using optimize de Brujin graph. We transform de Bruijn graph into series of equivalence sub graps. Euler path of every sub graphs represent the sub solution of the problem. Euler

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Semantic view of the proposed methodology for genome assembly; our architecture has two phases: graph generation and genome assembly. In graph generation phase, an optimized de brujin graph is generated and genome assembly phase different type sequence are assembled (i.e. Exon separation or Genome fragmentation etc.).



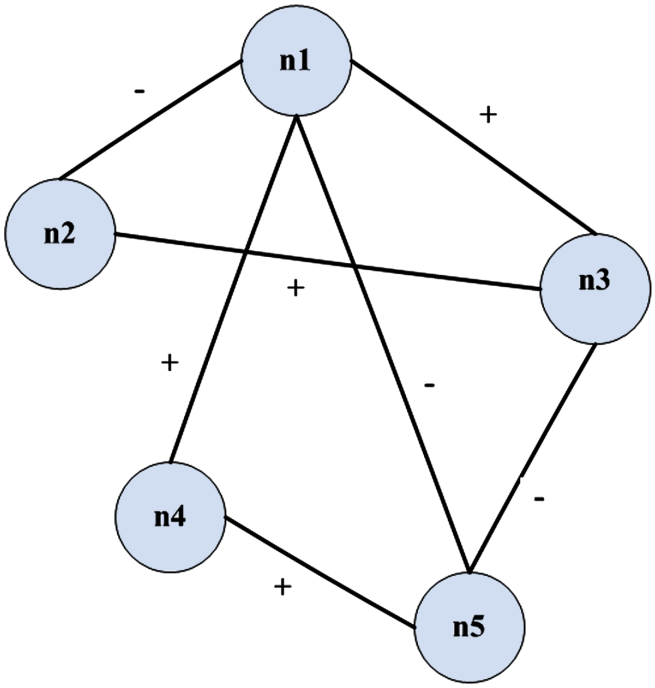
path is an efficient algorithm that is solved in a linear time. To combine the every solution of sub graphs represent the solution of gene annotation. In exon transcription, we marked initial and stopping sites in optimize de Brujin graph and find out euler path from intial site to stopping sites. Initial and stopping site indicates the exon annotation region.

SIGNED GRAPH

In this section we have introduced basic notations for signed graph. Let G=(V,E) be a finite undirected graph with multiple edges and self loop. The number of |υ| the vertices, is called the order of G and the connected number of |e| is called the degree of G. We write υ ϵ G if υ is vertex and e ϵ G if e is edge of G. The neighborhood of a vertex υ is $N_G(\upsilon)=\{u|(\upsilon,u)\ \epsilon\ G\}$.The vertex υ is isolated if $N_G(\upsilon)=0$.If a vertex has exactly one neighbor, it is denoted as a leap. We called G is discrete graph if all vertices are isolated. A subset A⊆G is stable if there is no edges (υ,u) with υ,u ϵ A. Graph A is complete if any two vertices of G are adjacent.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. A simple signed graph with positive or negative edges; an edge will negative if edges is connected with negative nodes and edges are positive if edges are connected with positive nodes.



A signed graph G=(V,E,ϕ) consists of vertices and edges (V,E) together with a labeling function ϕ:V→{+,-} of vertices V.A vertex υϵG said to be positive and negative if φ(υ)=+ and φ(υ)= - respectively. We let

$$G^{+} = \{\upsilon | \varphi(\upsilon) = +\} \text{ and } G^{-} = \{\upsilon | \varphi(\upsilon) = -\} \tag{1}$$

We say that a signed graph is negative if all its vertices are negative (Figure 3). Also, an edge $e=\{u,\upsilon\}$ is called negative, if $\upsilon,u \in G^{-}$.

If G and H constitute a two signed graph and its two disjoint vertex sets are V(G) and V(H), let $G \oplus H$ be their disjoint union, the vertex set of $G \oplus H$ is $V(G) \cup V(H)$ (Figure 4a) and its edge forms,

$$E(G \oplus H) = E(G) \cup E(H) \tag{2}$$

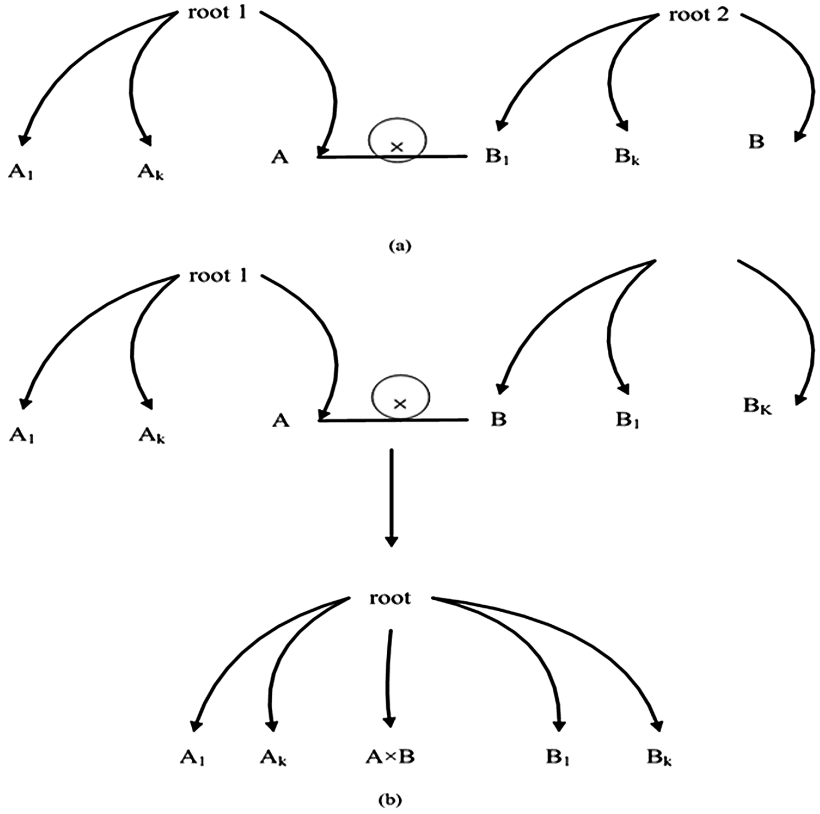
The complete connection $G \otimes H$ has the vertices set and the edge set is (Figure 4b)

$$E(G \otimes H) = E(G) \cup E(H) \cup \{(\upsilon, u) \mid \upsilon \in G, u \in H\} \tag{3}$$

We used undirected sign graph for graph analysis and transformation which are performed in graph reduction technique. It is more compatible and easy in graph analysis in which enlarging graph by joining new edges with information and reduction by deleting edges or replacing sub graphs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Two basic operations of signed graph: (a) And operation for two disjoint graph where two vertices convert into one vertex (b) complete operation for disjoint graph where two disjoint graph convert into one graph



GRAPH REDUCTION RULE

There are three basic fundamental operations of reductions rule for signed graph. The molecular operations are translated into following operations:

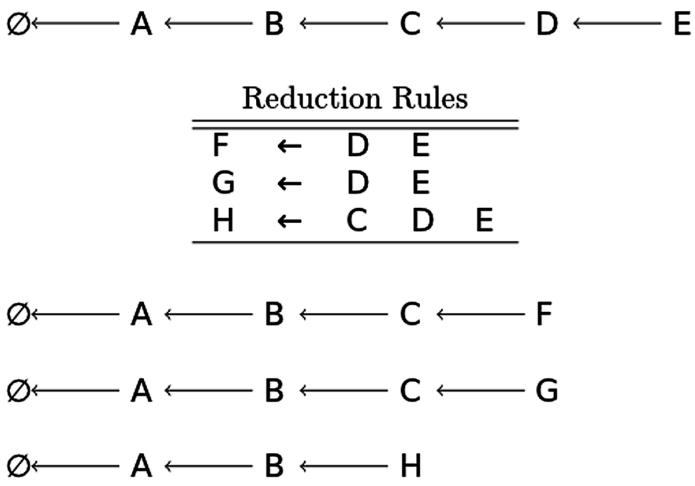
Let u and v be two vertices:

1. The negative graph rule for v is applicable to G if $\upsilon \in G^-$ and it is isolated in G. The result is the signed graph $nr_{\upsilon}(\upsilon)=G-\{\upsilon\}$. The number of vertices is $|nr_{\upsilon}|=\{\upsilon\}$.
2. The positive graph rule for v is applicable to G if $\upsilon \in G^+$. The result is the signed graph $nr_{\upsilon}(\upsilon)=G-\{\upsilon\}$. The number of vertices is $|nr_{\upsilon}|=\{\upsilon\}$.
3. The double rule for $v \neq u$ is applicable to G if $\upsilon, u \in G^-$ and $e=\{\upsilon,u\} \in E(G)$. The result for signed graph $dr(\upsilon) = G - \{(u,\upsilon), E', \varphi'\}$ where φ' is obtained to $G-\{u,\upsilon\}$ and E' obtained from complementary of E.

In basic reduction rule, graph rewriting technique is used for specification and generation of graph optimization. Graph analysis and transformation are performed by graph rewriting technique. Analyzing graph means enlarging graph by joining new edges with information and graph transformation means

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Graph reduction rule with termination point; in graph reduction rule it reduces the edges (paths) . A,B,C,D,E are nodes of the graph and Φ is termination point. D and E nodes are reduced by F or G.C,D and E nodes can also replaced by H. Here termination point indicates the ending point or pattern in transcription region. Replacement nodes reduce the graph complexity.



reduced into the graph rewriting by deleting and attaching sub graphs. In reduction rules we delete or replace two or more nodes by another node (Figure 5).

In graph reduction rules, node D and E are rewritten by F (Figure 5).Here D and E nodes can correlate with F, redundant nodes D and E are replaced by F. Node C,D and E are rewritten by new node H. Reduced edges are encrypted with new edge. Termination by edges accumulation and termination by edges subtraction is used for graph termination process. When, we reached to null point (Φ) that indicates the termination (Figure 5).

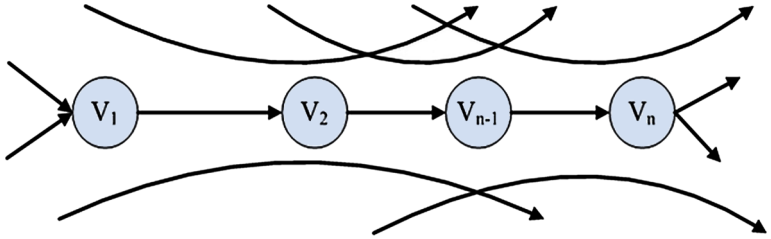
OPTIMIZE DE BRUJIN GRAPH

A sets of reads S={s_1…….s_n},define the de Bruijn graph G(S_1) with (l-1) vertices.An (l- 1) tuple v∈S_{l-1} is joined by directed with l-edges. If S_l contains l-tuple for which the first (l-1) nucleotides coincide with v and last (l-1) nucleotides coincide with tuple w. Each l-tuple from S_l corresponds to an edge in G. If S contains the only sequence S_1, then this sequence corresponds to a path visiting each edge of the de Bruijn graph. A de Bruijn can substitute every edge by k parallel edges, where k is the number of times the edge is used. If S contains the only sequence S_1,this operation creates k parallel edges for every l-tuple repeating k times in S_1.Euler path is an efficient algorithm that is solved in a linear time.

In de Bruijn graph, a vertex v is called a source if indegree(v) =0,a sink if outdegree(v)=0 and branching vertex if indegree(v).outdegree(v)>1.A path v_1.....v_n in the de Bruijn graph is called a repeat pattern if indegree(v_1)>1,outdegree(v_n)>1 and indegree(v_i)=outdegree(v_i)=1,for 1≤i≤n-1.Repeat pattern starts with v_i node and $v_{n,}$ are called exits from a repeat (Figure 6). An Eulerian path visits a repeat some times by visiting entrance and exit nodes. An Eulerian path covers a repeat if it contains an entrance into and exits from repeat by using end node. Every covering read-path reveals some information about the gene annotation between entrances and exit.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The overlapping pattern with repeated vertices and edges.v_1 ..... v_n are repeated vertices in an overlapping graph. It is difficult to find out the read -length and contigs. Euler path gives the optimal solution for overlapping pattern. This overlapped graph divided into two sub graphs.



To solve the Euler path for gene annotation, we have transformed the both graph G and path P into new graph G_1 with path P_1.This is called equivalence if it exits one to one correspondence in (G, P) and (G_1,P_1).We transform de Bruijn graph into series of equivalence transformation:

$(G,P) \rightarrow (G_1,P_1) \rightarrow (G_2,P_2) \rightarrow \ldots\ldots\ldots \rightarrow (G_k,P_k)$

To combine the every solution of sub graphs represent the solution of gene annotation (exon/introns separation), Euler path solution is used. We describe a simple equivalence transformation that solves the Euler path problem where graph G has no multiple edges. We consider two cases of transformation; one is x-y detachment and another one is x-cut. Let $x=(v_{in},v_{mid})$ and $y=(v_{mid},v_{out})$ are two consecutive edges of graph G and $P_{x,y}$ be the collection of all paths of P that includes all sub paths. $P_{\rightarrow x}$ defines as a collection of paths from P that end with x and $P_{y\rightarrow}$ as a collection of paths from P that starts with y. Then x,y-detachment is a transformation that adds a new edge $z=(v_{in},v_{out})$ and deletes the edges x and y from G (Figure 7a).This detachment transformation alters the systems of path P as follows:

1. Substitute z for x,y in all path from$P_{x,y}$.
2. Substitute z for x in all paths from $P_{\rightarrow x}$.
3. Substitute z for y in all paths from $P_{y\rightarrow}$.

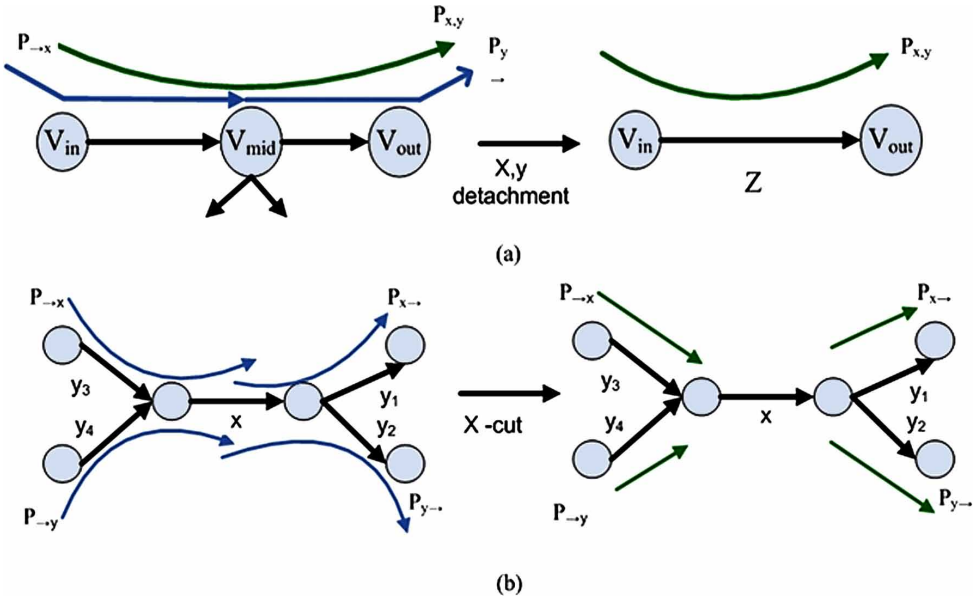
Every detachment reduces the number of edges in G and reduces the complexity of Euler path problem. Consider a fragment of graph G with 5 edges or 4 paths y_3-x,y_4-x,x-y_1 and x-y_2 (fig-7b).For graph fragmentation n edges or paths, it will be y_n-x,y_{n-1}-x,….y_1-x. In symmetric situation, x is tangle (repeated pattern) and there is no available information to relate any of paths y_3-x and y_4-x to relate another paths x-y_1 and x-y_2. An edge x=(v,w) is removable if

1. It is only incoming edge v and only one outgoing edge w.
2. x is either the initial or the terminal edge for every path P containing x.

An x-cut transformation P into a new system of paths by simply removing *x* from all paths in $P_{\rightarrow x}$ and $P_{x\rightarrow}$ without affected graph G itself (fig-7b).If x is a removable edge then x-cut is an equivalent transformation. Detachment and x cut proved to be powerful technique to build a simple de Bruijn graph and reduce fragments for all genome.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Equivalent transformation for Euler path; blue indicates the overlap paths and green line indicates equivalent paths.(a)X,Y –detachment which alters the systems of path for overlapped graph. In every detachment reduces the number of edges and reduces the complexity of Euler path. (b) X-cut which divide the overlapped graph into two or more sub graphs.



GENE ANNOTATION

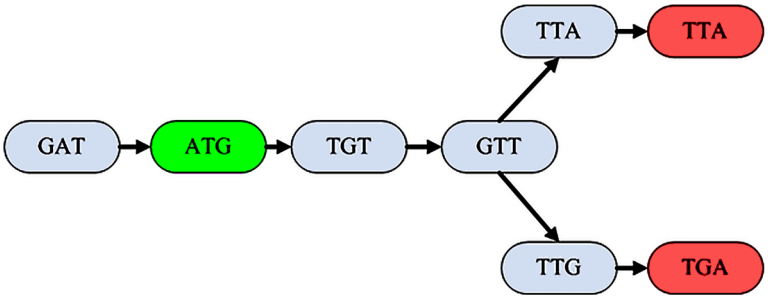
De Bruijn graph is used for gene Annotation (Exon Intron separation) and next generation sequence assembler. It reduces computational effort by breaking read (sort sequence) into smallest part of DNA. Reads are called k-mer where the parameter k denotes the length of bases for these sequences. De Bruijn graph capture exon separation with considering exon initial and stop sites using k-mer (Figure 8).Construct a de Bruijn graph for exon introns where separation consists of the following steps:

1. **Structure of k-Spectrum:** Reads are divided into overlapping sequence of k. Every k-mer consists transcription and stopping site in exon chain.
2. **Node Generation:** Every (k-1) node generates for k-spectrum. In de Bruijn graph, exon initial site is marked as initial node (source), ending node (sink) represents stop site of exon chain and intermediate node serves donor and acceptor site for exon transcription.
3. **Edge Construction:** A directed edge is created from x node to y node if there exists k-mer such that its prefix is equal to x node to y node. Overlap path are reduced to euler path simplification which described in section 4.1.To traverse marked source node to sink node we had indentified the exon chain from whole DNA sequence.

By reducing whole data set into k-mer overlaps the de Bruijn graph reduces the high redundancy in short read data set. In exon annotation, imitation sites start with ATG and termination site with TAA or TAG or TGA. Another donor and acceptor site generates the internal node in de Bruijn graph. Initial and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. The length of overlap k-1=2; green node indicates initial site and red node indicates termination site respectively. Black line indicates the order of k-mer. Two braches indicates the separate genome annotation and possible ways of exon chaining.



stopping site indicates the exon annotation region (Figure 8). By converting the set of reads into edges of the de Bruijn graphs, the annotation problem becomes equivalent to finding an Euler path graph. To reduce the exponential distinct Euler path, heuristics are usually applied to construct the graphs. The graphs are filtered of erroneous occurrences and nodes are unambiguously connected by edges which are merged together.

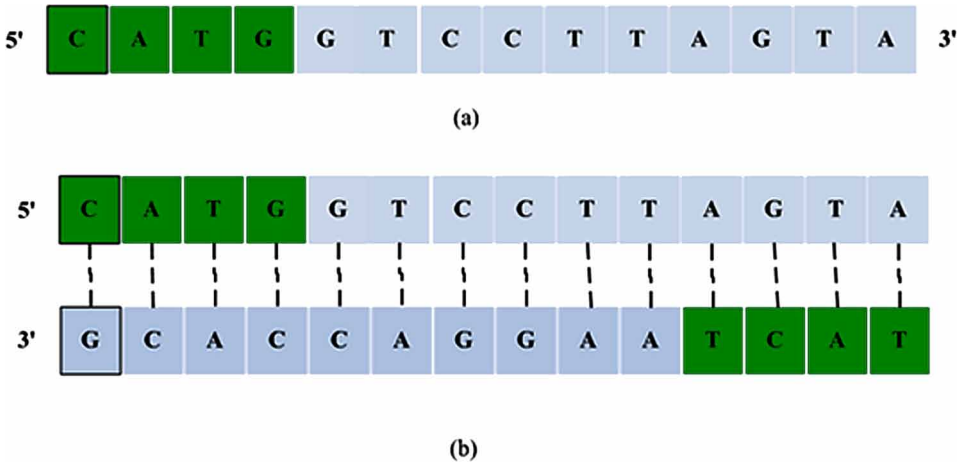
GENOME FRAGMENTATION

All sequencing platforms produce explanation of the target DNA molecule in the form of reads sequences of single-letter base pairs. (Miller et al., 2010). An example of a single-end read and a paired-end (Figure. 9a & Figure 9b). Genome assembly is two groups:

- **De Novo:** It is an assembly tools for DNA sequence. Contigs is A group of overlapping clones representing regions of the genome; the contiguous sequence of DNA created by assembling these overlapping chromosome fragments. In genome assembly scaffold is used to describe the first complete sequence assembly, to which new reads are aligned. Once have a scaffold you don't depend on de novo assembly anymore but can find a region to which any given new read is most similar and subsequently look for genetic variation.
- **Mapping/Reference:** In genomics complete genomes provide an outstanding resource for future sequencing projects. However, generating finished genomes is an exclusive and laborious endeavor. In some cases, the returns of finishing a genome are not worth the cost particularly when draft genomes can provide enough information for robust assembly process. Draft genomes contain a large number of unordered sequences of various lengths called contigs. Contigs can be joined into more contiguous sequences called scaffolds using paired-end reads. While a set of contigs/scaffolds can be useful for a wide variety of projects (gene annotation, gene expression profiling, etc.), contig order provides vital information in comparative genomics and analyses of specific genomic regions. Typically, a rough ordering of contigs can be accomplished by mapping contigs to a closely related reference genome.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Example of DNA fragment read (a) single end (b) double end; single end read indicates the single DNA sequence fragmentation and double end read fragment the multiple sequences. Double end read reduces the time complexity than single end.



As a result of the de novo assembly stage, a set of scaffolded contigs is accessible. Besides ordered contigs, scaffolds also contain gaps of estimated length separating the contigs. Scaffolds are also often called supercontigs. Low risk assemblies consistent with almost all of the detectable pair-wise read overlaps are called unitigs.Sequence alignment is the procedure of comparing two or more sequences by searching for a series of individual base pairs or base pairs patterns that are in the same order in all sequences. In most cases, the sequences cannot be aligned perfectly due to their mismatch. In an optimal case, mismatch base pairs and gaps are placed, in order to match together as many identical or similar base pairs as possible.

Sanger's method is most popular sequence fragment assembly algorithms were focused on creating multiple overlapping alignments of the reads to produce a layout assembly of the data. The alignment is used to read the consensus sequence, from which the DNA sequence is incidental. The described method was called the overlap-layout-consensus approach (OLC). OLC approach is one of the two typical approaches of the de novo DNA assembly, and is effectively used during the assembly of long, high-quality reads. Another very important, and currently most utilized approach in combination with Next Generation Sequencing (NGS) data, is based on the de Bruijn graph. This approach was commonly known as short read assemblers. Unlike the computationally expensive step of determining overlap amongst reads, the de Bruijn graph is constructed from fixed-length subsequences. If the sequence length is m then that have a perfect (m-1) suffix to (m 1)-prefix overlap. In de Bruijn assembler include Euler, AllPaths and Velvet methods.

Both the OLC and the de Bruijn approaches are graph based. Another type approaches also do exist. They include greedy search methods i.e. dynamic approaches such as Taipan (Schmidt et al., 2009).Errors in assembly can occur for two main reasons: Firstly, incorrect or incomplete information provided to the assembler. The real-world WGS data that contains errors can induce the following problems in overlap and de Bruijn graphs (Miller et al., 2010):

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Bubbles:** Bubbles are generated when two paths start and end nodes are same. Bubbles generated by causes of errors or biological variants. Dijksatra's and breath first algorithm is used the best path and remove such error.
- **Spurs:** Spurs are short, dead-end divergences from the main path. They are induced by sequencing error toward one end of a read. They can be induced by coverage dropping to zero. Possible causes include sequencing errors toward one end of a read, and low coverage.
- **Converging and Diverging Paths:** Reverse definition than for the bubbles, two paths converge into one, that later diverges again into two separate paths. Possible causes are repeats in the target or source genome.
- **Cycles:** Cycles are paths that converge on themselves. They are induced by repeats in the target. For instance, short tandem repeats induce small cycles.

Although sequencing technology has greatly enhanced, still no sequencing platform can produce sufficient data to assemble a complete genome from a single experiment. One of the key problems in shotgun sequencing is caused by repeats in genome sequences. Repeats, in cases when they are longer than fragment reads, persuade problems during the overlap phase of assembly. As a result, assemblies result in fragmented contigs, separated by gaps. Repeat caused fragmentation is more pronounced in the NGS technology. Assemblies containing gaps are called draft assemblies, while the process of filling the gaps is called finishing. Finishing involves obtaining missing sequences, improving low quality regions, resolving miss assemblies and ordering scaffolds. In order to fill gaps, further sequencing is usually required, either by amplifying and sequencing fragments spanning gaps, or by resequencing the entire DNA.

OVERLAP-LAYOUT-CONSENSUS APPROACHE

Let $N = \{n_1, n_2, \ldots\ldots n_m\}$ be a set of non-empty strings over an alphabet. An overlap graph of G is a complete weighted directed graph where each string in N is a vertex and the length of an edge between vertices x and y. The construction of the overlap graph is the first step of the OLC approach. General process of this approach consists of three main phases from which it derives its name:

1. **Overlap:** An overlap graph is created by joining all the reads by their respective best overlapping reads. This step is similar to the greedy approach.
2. **Layout:** The overlap graph is analyzed and simplified with the application of graph algorithms to identify the appropriate paths traversing through the graph and produce an approximate layout of the reads along the genome. The ultimate goal is a single path that traverses each node in the overlap graph exactly once.
3. **Consensus:** Multiple sequence alignment of all reads covering the genome is performed, and the original sequence of the genome being assembled is contingent through the consensus of the aligned reads.

A very important observation is that the identification of a path that traverses every node in the graph only once is a computationally difficult Hamiltonian path problem. The problem is also known as the Travelling Salesman Problem and is NP-complete.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The evolution of assembly Overlap-Layout-Consensus (OLC) and de-bruijn graph (DBG) algorithms are widely used OLC generally works in three steps: first overlaps (O) among all the reads are found, then it carries out a layout (L) of all the reads and overlaps information on a graph and finally the consensus (C) sequence is inferred. It is an intuitionist assembly algorithm. In the OLC algorithm, the identification of overlap between each pair of reads is explicit; typically by doing all-against-all pair-wise reads aligning. As a result, the OLC algorithm constructs a reads graph, which places reads as nodes and assigns a link between two nodes when these two reads overlap larger than a cutoff length. The nodes number is equal to the reads number, increasing linearly with sequencing depth and the links number will increase by a logarithmic scale.

DBG is an anti-intuition algorithm, working by first chopping reads into much shorter k-mers and then using all the k-mers to form a DBG and finally inferring the genome sequence on the DBG. The DBG assemblers were initially successful on small genomes such as bacteria, and were then extended to large genomes. In the DBG algorithm, the overlap relationship between neighboring k-mers is established implicitly. This process is completed by chopping all the reads into k-mers and simultaneously recording their neighboring relations. DBG algorithm constructs a k-mer graph that places k-mer as nodes and assigns a link between two nodes when these two k-mers are neighbors on the genome.

EXON TRANSCRIPTION

As exons are not independent, by splicing exons together to assemble a gene one can further eliminate false exon predictions by imposing translatability (i.e. adjacent exons must maintain the open reading frame). The main difficulty in exon assembly is the combinatorial explosion problem: the number of ways *N* candidates' exons may be combined grows exponentially with *N*. The key idea of computational feasibility comes from dynamic programming (DP), which allows finding "optimal assembly" quickly without having to enumerate all possibilities. We can limit possible errors by assuming each entry with a correct annotation that should satisfy:

- The initiation site is ATG.
- The donor site is GT.
- The acceptor site is AG.
- The stopping site is either TAA or TAG or TGA.
- No stop codon interrupts the open reading frames.
- The length of coding regions is a multiple of three.

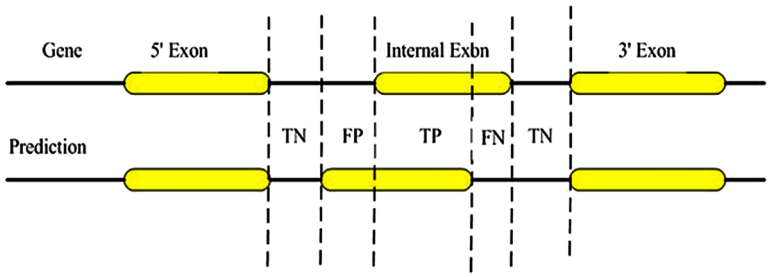
In pattern reorganization, the performance of a prediction system can be measured by the following statistics: true positive (TP), true negative (TN), false positive (FP) and false negative (FN). The internal exon prediction measurement on the nucleotide base pair in different level (Figure 10).

The accuracy of a prediction system is measured by sensitivity (SN), specificity (SP) and correlation co-efficiency (CC) as follows:

$$SN = \frac{TP}{TP + FN} \tag{4}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Measurement of Exon prediction comparing with predicted data set; True Positive (TP) indicates exon region consist both gene and predicted gene, True Negative (TN) indicates exon region not present both gene and predicted gene, in False Positive (FP) exon region is found in predicted pattern but not in real gene.



$$SP = \frac{TP}{TP + FP} \tag{5}$$

$$CC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN) \times (TN + FP) \times (TP + FP) \times (TF + FN)}} \tag{6}$$

RESULT

We experienced our gene analysis method on three data sets which are adh222, h178 and sag178 (Table 1). Adh22 is a single sequence of Drosophila melanogaster and 2.9Mb long. It is a well characterized sequence from the Adh22 region and has been used in the Genome Annotation Assessment Project(GASP). In the update version std1 contains 38 genes with a total of 111 exons and std3 contains 222 genes with a total of 909 exons. The genes lie on both strands. Both the authors of GENIE and of GENEID have used these two annotation sets for testing their programs. H178 is a set of 178 human genomic sequences with one complete gene each. Each contains one gene and a little neighboring DNA. The sequences are from EMBL, were compiled by[GAA+00] and have also been used by the author of GENSCAN for evaluation [YLB01]. The mean sequence length is 7169 bases. This set is known to contain three annotated genes which are very likely, if not known to be wrong. But we kept them for reasons of comparability. sag178 is a set of 43 sequences with 178 human genes. These sequences were also taken from [GAA+00]. The mean sequence length is 177 kilo bases (shortest 70, longest 282 kilo bases) and each sequence contains average 4.1 genes.

We measured the gene prediction accuracy with the sensitivity and specificity. A predicted exon is considered correct if both splice sites are at the annotated position of an exon. A predicted gene is considered correct if all exons predicted are correctly and no additional exons are predicted. Predicted partial genes were counted as predicted genes. For each data set these measures were computed globally. For comparing we choose GENESCAN and GENEID. We took GENSCAN as it is the most commonly used gene prediction program and as it is considered one of the best programs for humans. Our method considers memory efficient which is considering GENESCAN. We used GENEID (version 1.1) as there

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

is a special Drosophila parameter set available for it and as it uses a different approach not modeling the lengths. GENEID first finds splice site candidates, then exon candidates are using the splice site candidates and then genes using the exon candidates.

By combining both measures we use the f-measure, a good indicator of overall precision and recall of the algorithm. F-measure (Figure 11) indicates the positive predictive value of a data set.

$$F - measure = \frac{2 * SP * SN}{SP + SN} \tag{7}$$

Internal exon prediction is important for gene prediction. Internal exons also have flanking splicing boundaries: the acceptor splicing sites at the 5´ end and the donor sites at the 3´end. Potentially all of the selected donor site and acceptor site candidates can be paired to form exon boundaries. The number of internal exons in a gene is one less than the number of introns. In our learning data set, there are only 125 internal exons. When we paired 348 selected donor site candidates and 607 selected acceptor site candidates, we had got another set of 314 pseudo exons, The final internal exon prediction combined all the features of donor sites and acceptor sites, with the in-frame hexamer preferences and the log of exon length. The optimal discriminate function had the performance in Table 2.

Table 1. Accuracy result on training data sets

Data Sets	**Methods**				
			Euler	**GENESCAN**	**GENEID**
adh22	Base	SP	97.3%	96%	96%
		SN	91%	92%	92%
	Exon	SP	72.4%	71%	70%
		SN	61%	62%	57%
	Gene	SP	48%	47%	40%
		SN	34.7%	33%	29%
h178	Base	SP	96%	97%	89%
		SN	87%	86%	91%
	Exon	SP	82.7%	83%	66%
		SN	74.5%	75%	75%
	Gene	SP	38.5%	40%	14%
		SN	37%	36%	15%
sag178	Base	SP	93%	94%	89%
		SN	83.2%	64%	78%
	Exon	SP	67.4%	68%	67%
		SN	61%	45%	60%
	Gene	SP	20%	18%	17%
		SN	22%	14%	17%

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. F-.Measurement of gene prediction for different method; for different types of genome fragmentation Euler path based on de Brujin graphs gives the better result. Genemarker generates the more accurate result rather the GeneId algorithm.

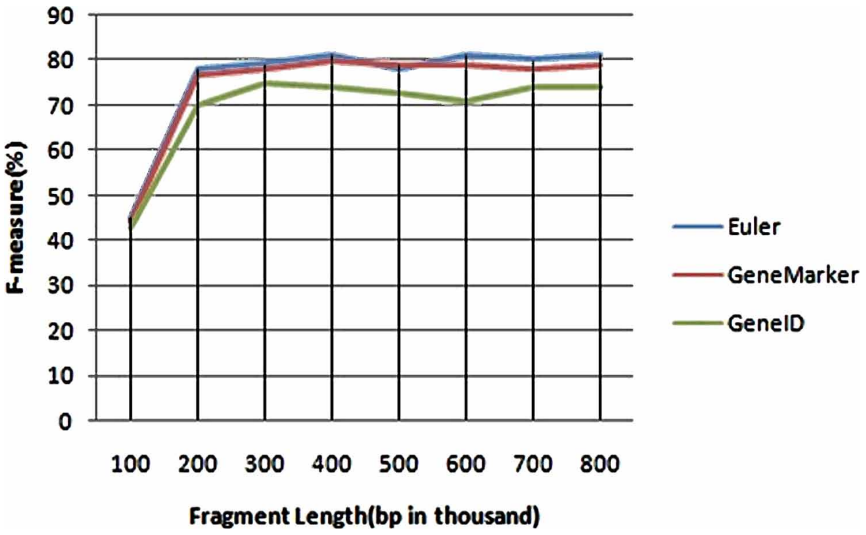


Table 2. Internal Exon measurement based on sensitivity, specificity, coefficient co-relation

Measurement	Total Number	Accuracy		
		SP	SN	CC
Base Pair	145	94.1%	87.2%	92.1%
Exon	325	91.0%	87.1%	87.1%

The Sensitivity and Specificity measure the estimate of the program capability of detecting reads that contains gene. Coefficient indicates the accuracy of gene detection. The Genome Analyzer can generate paired reads with a length of 150 bp and more. There are, however, alternate technologies such as sequencing that produce longer unpaired reads. In this experiment, we investigate the influence of read length on an assembly.We simulated paired reads of 150 bp length (with a 450 bp fragment size) and 400 bp unpaired reads from the E. coli genome and human chromosome 20, respectively. All reads were simulated as error-free and at 50× base coverage. We used Velvet with a k-mer size of 31 to assemble all four data sets.

The contig sizes for each assembly. The assemblies using the 150 bp paired reads yield far larger contigs as compared to the 450 bp single read assemblies (Table 3). This is in part due to the way a de Bruijn graph assembler such as Velvet works: the reads are split into smaller pieces, the k-mers. In this experiment, we used a k-mer size of 35 which is the maximum for Velvet . Therefore we compared in principle k-mers of length 35 with and without pairing information. The strong impact of read pairs as opposed to single ended reads is evident (Figure 12).

To examine how long inserts can improve the assembly we sequenced E. coli with different base pair size(6kb to 10 kb). All of the de novo assemblers that we tested have the ability to use long-insert size (Table 4). Table 4 shows results obtained using De Brujin graph with a variety of long insert libraries.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Effect of read length for different E.Coli sample with different base pair

Sample	N50 Contig Size	Largest Contig	Genome Coverage
E. coli, 150 bp pe	122,776 bp	326,886 bp	97.07%
E. coli, 450 bp sr	21,405 bp	127,976 bp	96.27%
Chr. 20, 150 bp pe	56,744 bp	401,312 bp	89.59%
Chr. 20, 450 bp sr	2,214 bp	23,423 bp	91.05%

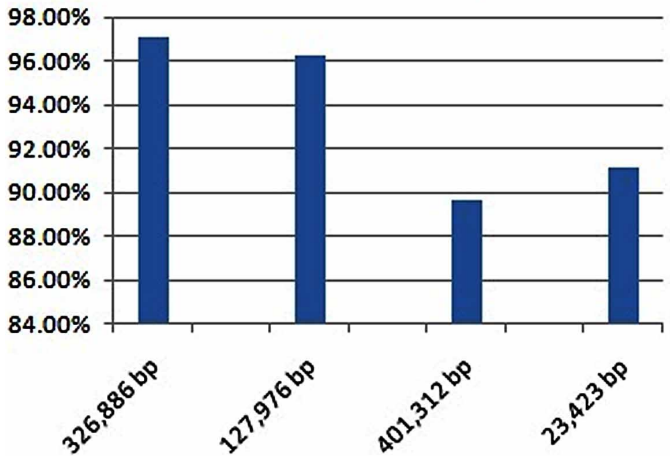
The 6 kb inserts already help bridge most of the repeats in E. coli if combined with the short insert sequence, and result in an assembly with an N50 of 1,403,330 bp. The two largest contigs cover a region of 2,348,095 bp, which corresponds to 68% of the E. coli genome. The rest of the genome is contained in smaller contigs (Figure 12).

The use of graph reduction rules to store a de Bruijn graph is straightforward and memory efficient. The expected false positive rate can be tuned based on desired memory usage, yielding a wide range of possible storage efficiencies . Optimized de bruijn graph is more efficient in time and space than the other assembler tools. We find out optimal euler path by using optimized de bruijn graph for gene annotation

Table 4. Effect of insert size for different super contig

Inserts	Reads (bp)	Coverage	N50 Super Contig	Largest Super Contig	Genome Coverage
200 bp	2×75	50×	94 kb	224 kb	98.88%
200 bp	2×75	50×	1.1 Mb	1.9 Mb	97.17%
200 bp	2×75	50×	4.2 Mb	4.2 Mb	98.59%

Figure 12. Genome accuracy analysis by using de brujin graph for four data set (E.coli, 150 bp pe, E. coli, 450 bp sr, Chr. 20, 150 bp pe, Chr. 20, 450 bp sr); we find better accuracy for E.coli, 150 bp.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

In this research, we have observed that graph theory gives better accuracy than other two models. Our method is a robust that continuously vacant the memory storage, especially the proposed algorithm truly stands out when the data set of DNA segments are increased. In fact, our simulation result indicates that is more accurate in large data set. It performs relatively well on the task of assembling exons to genes because programs with a similar exon level accuracy often have a lower gene level accuracy. This means those programs more often combine the exons to a wrong gene structure for example by splitting or joining genes. With the growing number of sequenced species, the possibilities of finding approximate possible exons and genome assembly by cross-species alignments of homologue genomic sequences also increase. This leaves the task of assembling possible exons or pattern to genes. In future, we shall consider this concept for finding possible introns analysis.

REFERENCES

Arthur, D., Kirsten, B., Edwin, P., & Steven, S. (2007). Identifying bacterial genes and endosymbiont dna with glimmer. *Bioinformatics (Oxford, England)*, *23*, 7. PMID:17237039

Besemer, J., & Borodovsky, M. (1999). Heuristic approach to deriving models for gene finding. *Nucleic Acids Research*, *27*(19), 3911–3920. doi:10.1093/nar/27.19.3911 PMID:10481031

Brown, T. A. (2002). *Genomes 2*. John Wiley & Sons Inc.

Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I. A., Belmonte, M. K., Lander, E. S., ... Jaffe, D. B. (2008). AllPaths: De novo assembly of whole-genome shotgun microreads. *Genome Research*, *18*(5), 810–820. doi:10.1101/gr.7337908 PMID:18340039

Chang, Z., Li, G., Li, J., Zhang, Y., Ashby, C., Liu, D., ... Huang, X. (2015). Bridger: A new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biology*, *16*(1), 30. doi:10.118613059-015-0596-2 PMID:25723335

Christopher, B. B. (1997). *Identification of genes in human genomicdna*. (Ph.D. Thesis). Stanford University, Stanford, CA.

Dalevi, D., Ivanova, N. N., Mavromatis, K., Hooper, S. D., Szeto, E., Hugenholtz, P., ... Markowitz, V. M. (2008). Annotation of metagenome short reads using proxygenes. *Bioinformatics (Oxford, England)*, *24*(16), i713. doi:10.1093/bioinformatics/btn276 PMID:18689842

Ehrenfeucht, A., Harju, T., Petre, I., Prescott, D. M., & Rozenberg, G. (2004). *Computation in Living Cells: Gene Assembly in Ciliates*. Springer. doi:10.1007/978-3-662-06371-2

Ewan, B., Michele, C., & Richard, D. (2004). *Gene wise and genome wise*. Genome Res.

Fryer, R. M., Randall, J., Yoshida, T., Hsiao, L., Blumenstock, J., Jensen, K. E., ... Gullansa, S. R. (2002). Global Analysis of Gene Expression:Methods, Interpretation, and Pitfalls. *Experimental Nephrology*, *10*(2), 64–74. doi:10.1159/000049901 PMID:11937753

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gens, P., Enrique, B., & Roderic, G. (2000). Geneid in drosophila. *Genome Research*, *10*(4), 511–515. doi:10.1101/gr.10.4.511 PMID:10779490

Green, P., Lipman, D., Hillier, L., Waterston, R., States, D., & Claverie, J. M. (1993). Ancient conserved regions in new gene sequences and the protein databases. *Science*, *259*(5102), 1711–1716. doi:10.1126cience.8456298 PMID:8456298

Greenblum, S., Turnbaugh, P. J., & Borenstein, E. (2011). Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(2), 594–599. doi:10.1073/pnas.1116053109 PMID:22184244

Handelsman, J. (2007). *Committee on Metagenomics*. Challenges and Functional Applications The National Academies Press.

Hoff, K. J., Lingner, T., Meinicke, P., & Tech, M. (2009). Orphelia:predicting genes in metagenomic sequencing reads. *Nucleic Acids Research*, *37*(Web Server), 101–105. doi:10.1093/nar/gkp327 PMID:19429689

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., ... White, O. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214. doi:10.1038/nature11234 PMID:22699609

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, *40*(D1), 109–114. doi:10.1093/nar/gkr988 PMID:22080510

Karim, R., Khaliluzzaman, M., & Mahmud, S. (2014). A Review of Image Analysis Techniques for Gene Spot Identification in cDNA Microarray Images. *International Journal of Artificial Intelligence & Application*, *5*, 1–7.

Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergstrom, G., & Behre, C. J. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature*, 4–10. PMID:23719380

Leila, T., Oliver, R., Saurabh, G., Alexander, S., Michael, B., Serafim, B., & Burkhard, M. (2003). Agenda: Homology-based gene prediction. *Bioinformatics (Oxford, England)*, *19*(12), 1575–1577. doi:10.1093/bioinformatics/btg181 PMID:12912840

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., ... Gordon, J. I. (2008). Evolution of mammals and their gut microbes. *Science*, *320*(5883), 1647–1651. doi:10.1126cience.1155725 PMID:18497261

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, *24*(3), 133–141. doi:10.1016/j.tig.2007.12.007 PMID:18262675

Margulies, M., Eghol, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., ... Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*, 376–378. PMID:16056220

Mario, S., & Stephan, W. (2003). Gene prediction with a hidden markov model and a new intron submodel. Oxford University Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Martin, G., David, K., Hari, T., & David, H. (2000). Genie–gene finding in drosophila melanogaster. *Genome Research*, 529–538.

Miller, J. R., Koren, S., & Sutton, G. (2010). Assembly algorithms for nextgeneration sequencing data. *Genomics*, *95*(6), 315–327. doi:10.1016/j.ygeno.2010.03.001 PMID:20211242

Mohammad, I. K., Sarwar, K., & Linkon, C. (2015). MSuPDA: A Memory Efficient Algorithm for Sequence Alignment. *Interdisciplinary Sciences Computational Life Sciences*, *6*, 1.

Morgan, X. C., Segata, N., & Huttenhower, C. (2012). Biodiversity and functional genomics in the human microbiome. *Trends in Genetics*, 1–8. PMID:23140990

Nimmy, S. F., & Kamal, M. S. (2015). Next generation sequencing under de novo genome assembly. *International Journal of Biomathematics*, *8*(05), 5. doi:10.1142/S1793524515300018

Nimmy, S. F., & Rahman, M. S. (2014). Algorithms for global protein–protein interaction network alignment. *Netw Model Anal Health Inform Bioinformatics*, *3*(1), 65. doi:10.100713721-014-0065-y

Noguchi, H., Park, J., & Takagi, T. (2006). MetaGene: Prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Research*, *34*(19), 5623–5630. doi:10.1093/nar/gkl723 PMID:17028096

Noguchi, H., Taniguchi, T., & Itoh, T. (2008). Meta Gene Annotator: Detecting Species- Specific Patterns of Ribosomal Binding Site for Precise Gene Prediction in Anonymous Prokaryotic and Phage Genomes. *DNA Research*, *15*(6), 387–396. doi:10.1093/dnares/dsn027 PMID:18940874

Peng, Y., Henry, C. M., Yiu, S. M., & Francis, Y. L. (2011). T-IDBA: A de novo iterative de Bruijn graph assembler for transcriptome. In Research in Computational Molecular Biology, Vancouver, Canada.

Powell, S., Szklarczyk, D., Trachana, K., Roth, A., Kuhn, M., Muller, J., ... Bork, P. (2012). eggnog v3.0: Orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Research*, *40*(D1), 284–289. doi:10.1093/nar/gkr1060 PMID:22096231

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., ... Wang, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, *464*(7285), 59–65. doi:10.1038/nature08821 PMID:20203603

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., ... Wang, J. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*, *490*(7418), 55–60. doi:10.1038/nature11450 PMID:23023125

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., ... Venter, J. C. (2007). The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biology*, *5*(3), e77. doi:10.1371/journal.pbio.0050077 PMID:17355176

Schmidt, B., Sinha, R., Beresford, B., & Puglisi, S. J. (2009). A fast hybrid short read fragment assembly algorithm. *Bioinformatics (Oxford, England)*, *25*(17), 2279–2280. doi:10.1093/bioinformatics/btp374 PMID:19535537

Simpson, J.T., Wong, K., Jackman, K., Schein, J.E., Jones, S.J., & Birol, I. (2009), A parallel assembler for short read sequence data, Res. ABySS.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., ... Natale, D. A. (2003). The COG database: An updated version includes eukaryotes. *BMC Bioinformatics*, *4*(1), 41. doi:10.1186/1471-2105-4-41 PMID:12969510

Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., ... Banfield, J. F. (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*, *428*(6978), 37–43. doi:10.1038/nature02340 PMID:14961025

Wang, S., Cho, H., Zhai, C. X., Berger, B., & Peng, J. (2015). Exploiting ontology graph for predicting sparsely annotated gene function. *Bioinformatics (Oxford, England)*, *31*(12), i357–i364. doi:10.1093/bioinformatics/btv260 PMID:26072504

Wilke, A., Harrison, T., Wilkening, J., Field, D., Glass, E. M., Kyrpides, N., ... Meyer, F. (2012). The M5nr: A novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics*, *13*(1), 141. doi:10.1186/1471-2105-13-141 PMID:22720753

Wu, D., Wu, M., Halpern, A., Rusch, D. B., Yooseph, S., Frazier, M., ... Eisen, J. A. (2011). Stalking the fourth domain in metagenomic data: Searching for, discovering, and interpreting novel, deep branches in marker gene phylogenetic trees. *PLoS ONE*, *6*(3), e18011. doi:10.1371/journal.pone.0018011 PMID:21437252

Xu, L., Chen, H., Hu, X., Zhang, R., Zhang, Z., & Luo, Z. W. (2006). Average Gene Length Is Highly Conserved in Prokaryotes and Eukaryotes and Diverges Only Between the Two Kingdoms. *Molecular Biology and Evolution*, *23*(6), 1107–1108. doi:10.1093/molbev/msk019 PMID:16611645

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., & Dominguez, M. G. (2012). Human gut microbiome viewed across age and geography. *Nature*, 10. PMID:22699611

W.Yinan, D. W. Renner, I. Albert, & M. L.SzparaL.(2015).VirAmp: a galaxy-based viral genome assembly pipeline GigaScience 4,19.

Yuzhen, Y., & Haixu, T. (2015). Utilizing de Bruijn graph of metagenome assembly for metatranscriptome analysis. *Bioinformatics (Oxford, England)*, 1–8.

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*(5), 821–829. doi:10.1101/gr.074492.107 PMID:18349386

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, *7*(1-2), 203–214. doi:10.1089/10665270050081478 PMID:10890397

KEY TERMS AND DEFINITIONS

De Brujin Graph: A generalized de Bruijn graph is a directed graph, where an edge exists from all nodes of form $V=(v_1, v_2 \ldots\ldots v_n)$ to $U=(u_1 u_2 \ldots\ldots u_n)$ with U being a left-shifted version of V, or $u_1=v_2$, $u_2=v_3$,.....$u_{n-1}=v_n$.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DNA Fragmentation: DNA fragmentation is the breaking of DNA molecules into smaller pieces. Fragmentation of DNA is an early step in next generation sequencing workflows, as well in the creation of DNA inserts for expression libraries.

Euler Path and Circuit: An Euler path is a path that uses every edge of a graph exactly once. An Euler circuit is a circuit that uses every edge of a graph exactly once path.

Exons: A segment of a DNA or RNA molecule containing information coding for a protein or peptide sequence.

Gene: A linear sequence of nucleotides along a segment of DNA that provides the coded instructions for synthesis of RNA which, when translated into protein, leads to the expression of hereditary character.

Gene Annotation: Annotation is the process by which relevant information about these raw DNA sequences is added to the genome databases. This involves describing different regions of the code and identifying which regions can be called genes.

Gene Assembly: Genome assembly algorithms are sets of well defined procedures for reconstructing DNA sequences from large numbers of shorter DNA sequence fragments.

Genome: The total amount of genetic information in the chromosomes of an organism, including its genes and DNA sequences.

Introns: A segment of a DNA or RNA molecule that does not code for proteins and interrupts the sequence of genes.

Metagenomic Gene: Metagenomics is the study of a collection of genetic material (genomes) from a mixed community of organisms. Metagenomics usually refers to the study of microbial communities.

Signed Graph: A signed graph is a graph in which each edge has a positive or negative sign. Formally, a signed graph Σ is a pair (G, s) that consists of a graph $G = (V, E)$ and a sign mapping or signature s from E to the sign group $\{+,-\}$.

This research was previously published in Classification and Clustering in Biomedical Signal Processing edited by Nilanjan Dey and Amira Ashour; pages 290-314; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 60
Healthy 3.0:
Healthcare Digital Dimensions

Maria Teresa Borges Tiago
University of Azores, Portugal

Flávio Tiago
UAC, Portugal

Francisco Emanuel Batista Amaral
University of Azores, Portugal

Sandra Silva
Unidade de Saúde de Ilha de S. Miguel Azores, Portugal

ABSTRACT

Over the past few years, healthcare practice has evolved to include new forms of information and communication technologies, and healthcare providers have begun to leverage IT solutions and the internet to reach consumers in transformative ways. There is a common thought both in business and academia that the technology adoption process is a key component of success and allows firms to achieve and sustain competitive advantages. Therefore, this research attempts to reinforce the assessment of ICT impacts on healthcare, analyzing it from two different perspectives: 1) firm performance, which is measured as a combination of induced and intangible benefits beyond the traditional financial benefits; and 2) healthcare providers' acceptance and adoption of ICT tools. To assess the first perspective, a structural equation model was applied to a large database sample covering firms from 17 European countries. The results reinforce the importance of induced and structural benefits in firms' overall performance. The second perspective was analyzed only for the Portuguese healthcare providers sample. These results can be a starting point for rethinking the healthcare models emphasizing the perspectives of Healthy 2.0 and considering that patients' technological pattern evolution will lead to Healthy 3.0 in the short term. Nevertheless, some questions remain unanswered regarding the impacts of the ICT acceptance process on overall benefits; therefore, future research will focus on this domain.

DOI: 10.4018/978-1-7998-1204-3.ch060

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

It is no longer enough to consider the impact and role of information and communication technologies (ICTs) at firm and industry levels. Much has been written about these matters (Inklaar, O'Mahony, & Timmer, 2005; Jennings, Clifford, Fox, O'Connell, & Gardner, 2015; Pilat, 2004; M. B. Tiago & Tiago, 2013; Van Ark, Inklaar, & McGuckin, 2003), but the vast technological changes that occurred require that the scope of these researches be broadened.

Traditionally, and taking into account the results found in previous studies, the impact of ICT can be perceived as detailed in: 1) the contribution of ICT investment to capital intensity; 2) the contribution of the technological adoption model to a firm's performance; 3) the spillovers of rapid innovation by the use of ICT; 4) the competitive and dynamic effects of ICT on a firm's turbulence in terms of entry, exit, and the mobility of market shares of incumbents; and 5) the empowerment of strategic dimensions through the use of ICT tools (Fracolli et al., 2014). Nowadays, there needs to be the added consideration of the role and impact of ICT in the communication flows of the firms, regardless of internal or external context. In these contexts, the web developments influenced the flows of information creation and sharing. Additionally, the initial Web evolved into Web 2.0, reflecting a gradual growth in terms of popularity and resources and the emergence of new tools that meet users' need for interactivity with information (Tekli et al., 2013).

ICT-related sectors are often characterized as high-tech and knowledge-intensive industries, since they rely deeply on the use of ICT tools to leverage an internal innovation process When approaching other service sectors, the impacts of ICT adoption are perceived as lower (I. Chang, 2007). There has been much discussion of the role that recent advances in ICTs could play in improving health service systems in developing countries, as well as other services industries with traditionally lower levels of technology connected to healthcare (Hadgkiss & Renzaho, 2014).

Drawing on the existing literature, this chapter attempts to analyze: 1) the impacts of ICT in healthcare services firms' performance; 2) the adoption and acceptance of ICT patterns of physicians and nurses; and, 3) to unveil some of the trends in the healthcare sector driven from ICT use in European countries. Therefore, a conceptual model was developed for understanding and investigating the different contributions achieved by the adoption of ICT in this sector.

The data used in this research came from the 2006 e-business watch survey, covering 834 health services firms in 17 countries (Belgium, Bulgaria, Czech Republic, Denmark, Finland, Greece, Spain, France, Hungary, Ireland, Italy, Latvia, The Netherlands, Norway, Poland, Sweden, and the UK). The survey was enlarged with data gathered in Portugal regarding physicians' and nurses' use of technology-based systems.

In general, our findings support the conceptual framework presented and the results obtained can be used as a support tool in ICT decision makers' processes, since it clearly states the overall benefits of ICT successful implementation as well as presents the global performance impacts achieved with an integrated IS framework. Some useful preliminary insights are produced, however, leaving a considerable number of issues for future research, including the possibility of extending the research to include new tools and perspectives.

This chapter is organized in the following manner. The next section summarizes the literature related to the adoption of ICT tools by healthcare firms. The third section of the chapter formalizes the major points presented in the conceptual model and translates them into hypotheses. In the fourth section,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the methodology is applied and the results are presented. The fifth section offers concluding remarks about the some of the major issues, controversies, and problems, as well as suggests solutions and future research paths.

BACKGROUND

In the mid-1980s, information and communication technologies have achieved a dominant role, leading to several changes in the internal organizational environment, regardless of the sector. In today's competitive global context, such technologies act as an imperative in two ways: 1) reducing the uncertainties related to production and services processes (Henderson et al., 2012), as well as supporting administrative tasks (Dewett & Jones, 2001; Sheng, Chang, Teo, & Lin, 2013) and subsequent achievement of sustainable competitive advantage (McKee, Varadarajan, & Pride, 1989); and 2) narrowing the distance between providers and customers, allowing a continuum of information flow that enhances overall service quality (Darkins, Foster, Anderson, Goldschmidt, & Selvin, 2013; Egea, González, & Menéndez, 2015; Kumar, Novak, & Tomkins, 2010; McGowan et al., 2012).

These changes impel rethinking firms' strategies concerning the digital domain: as IT investments that need to be performed to enhance performance; and, as a daily tool between service providers and clients that needs to be spread out and regulated.

A review of the existing literature related to information and communication technologies and IS impacts on organizational performance may enable a more effective implementation of these tools and how organizations can take advantage of their application (Bayo-Moriones, Billón, & Lera-López, 2013; Eder & Igbaria, 2001). According to the results of several research studies, such as the papers of Papastathopoulos, et al. (2007) and Peansupap and Walker (2005), there are two main reasons that tend to influence greatly the way firms handle the acceptance and implementation of ICTs: the level of investment involved and failure rates or past experiences of failure associated with the implementation of ICT applications.

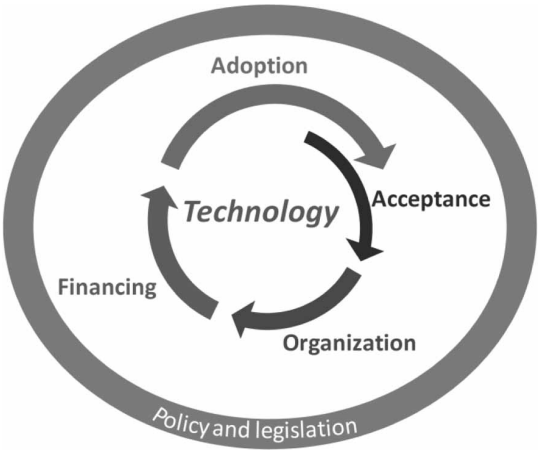
More, the debate regarding the use of ICT seeks to determine which firms' characteristics enhance the adoption of ICT and enhance their impact on organizational performance (Dewett & Jones, 2001; Liao, 2003). In this sense, the literature suggests as adoption enablers' firms' characteristics, strategy commitment and external environment. This concept has stimulated the study of the elements and drivers of adoption of ICT/IS use in recent years (Bayo-Moriones et al., 2013; Bayo-Moriones & Lera-López, 2007; Liao, 2003; Liu & Liu, 2008; Papastathopoulou et al., 2007; Schillewaert, Ahearne, Frambach, & Moenaert, 2005; Wild & Griggs, 2008; Yi & Hwang, 2003 ; Zhu & Kraemer, 2002; Zhu, Kraemer, & Dedrick, 2004).

Information technology has become a common asset in health services industry in the past four decades and has gained widespread usage (Rahimi & Vimarlund, 2007). Thus, it's hard to picture health care services with absence of information and communication technology tools (I. Chang, 2007; M.-Y. Chang, Pang, Tarn, Liu, & Yen, 2015; Hung, Hung, Tsai, & Jiang, 2010). Nevertheless, the evaluation of the outcomes, the benefits and inhibitors associated with IT-based systems in medical informatics is still a challenge.

Nowadays, it is clear that ICT offers tremendous opportunities to improve daily-base activities in healthcare sector. One of the main reasons for the growing importance of IT in this sector relies on the fact that IT is consider to be an enabler to handle the enormous amount of information related to patients.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. ICT adoption dimensions in healthcare sector



The identified determinants, which would influence the future implementation of electronic healthcare systems, can be classified into six major categories: (1) Technology, (2) Acceptance, (3) Organization, (4) Financing; (5) Adoption; and, (6) Policy and Legislation (see Figure 1).

The use of *technology* in health services can be observed in regular operations such as electronic patient records access, decision support systems which helps clinical actions, among others (Ammenwerth, Gräber, Herrmann, Bürkle, & König, 2003; Buntin, Burke, Hoaglin, & Blumenthal, 2011; Egea et al., 2015). The following table presents some of these applications and the research made around them (see table 1).

After summarizing all these works that analyze the use of digital tools on healthcare, it seems very cliché to define the concept of healthcare, but the panoply of definitions and uses found in literature impels us to do so. Therefore, in this work "Health care" is considered as the broad concept, which encompasses diagnosis and treatment of patients, education of staff, patients and the general public, and administrative activities (Fischer & Soyez, 2015).

From the results of multiple studies, it can be seen that information technologies such as electronic patient records, electronic health records, e-prescribing, decision support systems, electronic management of chronic disease, and bar coding of drugs and biological products can reduce health care costs and medical errors (Ammenwerth, Schnell-Inderst, Machan, & Siebert, 2008; Bates et al., 1998), as well as improve the communication between all stakeholders (Lau, 2011; McGowan et al., 2012). However, implementation of these information technologies has lagged in most European nations as well as the United States, leaving a question unanswered: *if the adoption of ICT has so many positive contributions, why it's use still so reduce and controversial?*

The quest for revealing this answer leads us to the second dimension – *acceptance* – and to its four different domains: social and ethical barriers; interoperability of systems; technology organizational environment and end-users acceptance.

ICT adoption and especially internet use is rapidly changing the landscape of health information (Ahmad et al., 2006; Sharma, Kilian, & Leung, 2014). Nevertheless, the potential arise from this electronic elements is far bigger then what is been explore. The full capacity use is still constrained by social and ethical barriers.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Examples of research performed on technology tools adopt in healthcare services

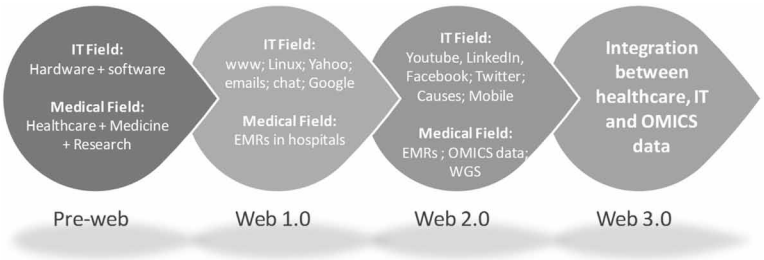
Authors (Year)	ICT Tool Analyzed
Laerum et al (2003); Stengel et al (2004)	Computer-based documentation
Wang (2003); Adams et al (2003); Hier et al (2004); Likourezos et al (2004)	Electronic medical record
Hippisley-Cox, et al. (2003); Mäkelä et al (2005)	Electronic patient record
Ammewerth et al (2003)	Computer-based nursing process
Baum et al (2004)	Helpdesk in HIS
Laerum et al (2004)	Hospital information system
Scott et al (2005)	Electronic health record
Persaud et al (2005); Steventon et al. (2012); Darkins et al. (2013); Henderson et al. (2012)	Telehealth
Noble et al (2005)	Telemedicine
Ahmad, Hudak, Bercovitz, Hollenberg, & Levinson (2006)	Electronic consultation
Buntin, Burke, Hoaglin, & Blumenthal (2011); Rahimi and Vimarlund (2007)	Health information system
Lau (2011)	Web 2.0 Tools
McGowan et al (2012); Fischer and Soyez (2015)	Social media applications
Costa Fabricio F Costa (2013)	PatientsLikeMe social network
Egea et al. (2015)	Adoption of Internet technologies by physicians

The work of Fabricio F. Costa (2013) added to this field by incorporating the technology as a catalyst and enhancer of healthcare over time (see Figure 2).

Since web 2.0, big data start to be used and in these processes the primary stakeholders in the decision of, and the impact of, are the patients, the researchers, and the other healthcare providers. Additionally, it can be considered in the participations of payers and other regulators. With web 3.0, genetic, genomic and other "OMICS" information tends to be integrated to help services providers to deliver a person-driven healthcare.

Thus, as mention by Jensen et al. (2009) and reinforce by Fabricio F. Costa (2013) and Brothers and Rothstein (2015), clients' questions regarding data privacy and security led to smaller rates of implementation of electronic health systems. By one hand, the legal frames in different countries constrain

Figure 2. IT and Medical evolution
Source: Adapted from Fabricio F. Costa (2013)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the use and electronic diffusion of patients' records. By other hand the patient and the physician have ethical and personal concerns regarding the adoption of these tools, especially when these enable the access by third parts (Brothers & Rothstein, 2015).

Lucas (2008) examining the adoption of ICT in developing countries uncover some of the potential problems related to the technology environment of each business unit. To this author in the short term is quite easy to have goods results driven from the adoption of ICT. Nevertheless, this initial small-scale success aren't guarantee of sustainable large-scale benefits or even long term success. The main reasons for this septic framework are related to robustness of sophisticated electronic equipment; the availability of technical staff to operate, maintain and support the equipment adopt; the complexity of hardware and software failures; the adequate training programs; the proper funding to finance routine operating expenses, maintenance costs and the frequent needs for upgrades; and the end-users commitment and knowledge regarding the system benefits and applications.

The second dimension is related to interoperability of the information systems, since healthcare providers use information systems with several different types of data and user interfaces (Macedo & Isaías, 2009) and standardized output is need. Until recently, IT products available for healthcare providers were mostly designed for large organizations and were costly. And, as Brannigan (1985) suggested, the major problem with IT acquisition is not the initial decision process, but the complete final analyses of adequacy of the system to the health care providers needs. Technical verification and validation are important keys during system development, but their validation after implementation is critical.

The fast and growing adoption, combine with the need to evaluate the outcome of the investment, include the "technology assessment" as a crucial element of health policy (Fischer & Soyez, 2015). Thus, back in 1990, in a conference held in Montpellier, Flagle et al. (1991) stated out the importance of (1) ''develop a dialogue between the fields of medical informatics and Health Technology Assessment''; and (2) ''to issuing recommendations about the methodological requirements for evaluating Health Information Systems''.

As research in other fields suggests, the adoption and acceptance of ICT tools requires a set of human, organizational and technological conditions. Since the seventies, research on information systems has contribute to a better understanding of the advantages related to the use of IS and the facilitators factors of an ICT acceptance process (Hadgkiss & Renzaho, 2014; Pilat, 2004). The work of Davies (1989) brought some light to his matter, introducing a new framework of analysis – Technological Acceptance Model and since then several studies have been conducted analyzing the adoption of technologies by end users (T. C. Davis et al., 1993). On his model, Davis examines the mediating role of perceived ease of use and perceived usefulness in their relation between systems characteristics (external variables) and the probability of system use (an indicator of system success). An update version of his research -TAM2- has presented several years later and the overall results explain almost 40% of systems use. In both works the emphasis has been set on end users performance. This model has been test in healthcare sector (M.-Y. Chang et al., 2015; Hu, Chau, Sheng, & Tam, 1999; Raitoharju, 2005) and the results pointed to the fact that individual impacts are often neglected when evaluating ICT initiatives as well as the IT stress as a factor of IT acceptance.

Another approach is to analyze the customization of the systems to healthcare providers' main characteristics. Considering the strong push for clinical leadership in the development and procurement of information technology in health care, it's surprising the lack of clinical input acknowledge by IT system' developers (Heathfield, Pitty, & Hanka, 1998).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The lack of participation and commitment to the system implementation and outcome continuous to be consider the major factor in the failure of information technology in health services (Lucas, 2008). More recent work has built on the practice perspective recognizing the importance of contextual environment in ICT adoption. Existing work practices may shape and be shaped by the introduction of new ICTs (Egea et al., 2015; Hadgkiss & Renzaho, 2014).

Back in the earliest nineties, Tornatzky and Fleischer (1990) present a framework which assess technological and organizational environment through three key dimensions - the organizational context, technological context and the surrounding context. Since then, the TOE model has been widely used, since it helps to understand the technological path of the firm.

Papastathopoulou (2007) and Peansupap and Walker (2005), present two other strong reasons that tend to influence greatly the way firms handle the acceptance and implementation of ICTs: the level of investment involved and failure rates or past experiences of failure associated with the implementation of ICT applications. These ideas introduce us in to the fourth dimension: *financing*.

Despise the reports revealing that electronic medical records provide financial benefits by helping to reduce costs and improve revenues, few formal cost-benefit analyses have been performance (Jennings et al., 2015). Wang et al (2003) analyzed the results obtain by organizations during a 5 year period driven from the adoption of ICT. According to these authors the healthcare firm had two majors costs associated with this implementation: systems cost, which are related to system implementation, support and maintenance; and induce cost, which are related to the process of converting paper-base operations in digital operations. These two domains had an overall cost of almost $50.000, which was lower than the benefits achieved of $86.400. The benefits obtained in the process were found to be especially savings in drug expenditure, reduction of cost in radiology and decreasing of billing expenditure costs. Similar conclusions were found in the work of Ruland and Ravn (2003), Hillestad et al. (2005) and Lapointe and Rivard (2006). However, traditional financial and accounting methods of evaluation are not totally suitable in this case, because there are some intangible, indirect and even strategic benefits related to IT implementation that need to be considered (Grembergen & Amelinckx, 2002).

Looking at the healthcare sector, back in the 70s, among others authors, Ware, Davies and Stewart (1977) suggested the increasing importance of customer/patient satisfaction with medical and healthcare services. Almost four decades after those statements, patients' satisfaction remains a critical aspect; however, as services literature notes, the need for employees' satisfaction is a driver for customers' satisfaction, since the healthcare service outcome is dependent on their performance (Anderson, Fornell, & Rust, 1997). Several studies conducted to measure patients' overall satisfaction confirm the importance of physicians and nurses in this process (Jennings et al., 2015).

The fifth dimension comprises the *adoption* of different healthcare information technologies to improve the quality of healthcare services. In this sense, improving the quality of care and customer satisfaction are key factors that need to be considered and that can push healthcare providers to adopt more ICT tools (Hung et al., 2010). These tools can improve services quality, since it changes the model of health services delivery (Jennings et al., 2015). Some examples of such services include educating patients about their diseases, online follow-ups, monitoring of symptoms, and clarifying doubts about diseases. To support these new services, adoption healthcare providers can exploit different IT applications.

Innovative companies found that Web 2.0 collaborative applications could be introduced in their current or new implementations of intranets, bringing them advantages in internal communication with employees. Given the characteristics of blogs, RSS and wikis, they are the most used Web 2.0 tools to facilitate communication, collaboration and sharing of internal information.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Several studies have examined the Web 2.0 intranets tools (Lai & Turban, 2008; Richter, Stocker, Müller, & Avram, 2013) and the outcomes of their use:

- Wikis: Collaboration, audit, internal training, knowledge management;
- Blogs: in-house training, communication, sharing, definition of social responsibility, knowledge management;
- RSS: News, information and knowledge management.

At a macro level, these information flows can promote inward sharing, outward sharing, inbound sharing and outbound sharing (Drapeau & Wells, 2009; Murphy, 2010). Looking at the inward sharing in healthcare organizations, the mere fact that providers belong to small groups who work together within the organization improves their subjective wellbeing, and allows them to share knowledge easily between colleagues to gain expert status, to collect feedback and to improve overall job satisfaction (Murphy, 2010).

The outbound level and the outward sharing have gained relevance in the past decade with the explosive growth of the Internet, which stimulated social behaviour exchanges and consequently modified human daily activities, habits and interactions. Online communities explore the ubiquity of the web and bring people together from all over the globe, allowing them to share knowledge and promote a collective dialog (Budden, Anthony, Budden, & Jones, 2011; Kumar et al., 2010).

In this context, healthcare institutions were forced to adopt a user-centric model that doesn't focus on diseases, but on the user as an involved member of the decision-making process, with capabilities to choose between different health options and preventive health behaviors. As a result, the "P4 Medicine" model seeks to change the medicine perspective from reactive to proactive, aiming to maximize individual health rather than just treat the causes of disease. The P4 in this model stands for: predictive, preventive, personalized, and participatory (Fabricio F Costa, 2013).

On the second and fourth P, the use of social networks leverages the doctor/patient interaction, allowing the sharing of information among patients as the network (Fabricio F Costa, 2013). A good example is the PatientsLikeMe social network.

The potential of social networks and other Web 2.0 tools to establish new health services cannot be ignored. These services are known as Healthy 2.0, which reflects the use of social media by health professionals and patients in order to customize treatments and promote health education (Sharma et al., 2014).

Several studies analyzed the adoption of social networking to mediate the communication flows between organizations and patients, and unveiled the most common applications: health information and dissemination campaigns, and tele-assistance and monitoring. But these studies also revealed that social networks are been used by patients as a community in order to share and search information and experiences with peers, to keep and search medication information, and to use an interactive medication database (Fabricio F Costa, 2013; Dhillon, Wünsche, & Lutteroth, 2013; Sharma et al., 2014).

More, this approach leads to the use of OMICS and other big data driven-content, but also requires a human-IT approach that enhances the adoption and use of technology by all stakeholders (Fischer & Soyez, 2015).

So, the question that remains unanswered is: if technology enables healthcare providers' to deliver a high quality service, why some many are reluctant to engage in a more digital service? This question finds a partial answer in the legal domain, since some of the inhibitors found in the literature are related with legal risk and privacy concerns (Classen et al., 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The last dimension cover refers to *policy and legislation* and finish what can be consider as the cycle of adoption of ICT in healthcare (see, Figure 1). In this domain there are no general rules and guides, since it's deeply associated with national health systems policy and they differ from one country to another. Nevertheless, there are some common points that need to be considered: patients' privacy rights and their legal protection (Jensen et al., 2009); security standards which allow a common language between players (Gritzalis, 1997); and, clinicians activity rules and structure (Buntin et al., 2011).

The Council of European Professional Informatics Societies, back in 2005, produce a report considering the Standardization for ICT Security and one of the sections was devoted to analyze the ICT security standards for healthcare applications. The authors advocate that Health informatics standards were essential to achieve the goals of e-Health in Europe, supporting interoperability between systems and patient information exchange between health organizations and protecting patients and clinicians' privacy rights. Pirnejad et al. (2007) suggested that the first steps in establishing an international application with adequate data protection safeguards rely in building an Inter-organizational communication network (IOCN) in healthcare with appropriated strategies safeguarding security measures. For instance, in the European case the directive "CEN EN14484 (2003) International transfer of Personal health data" covered data protection and established specific terms of exchange data with third countries.

Although healthcare researchers and information system investigators have studied ICT adoption to some extent, there are still few conceptual bases available to assess the global information and communication technology contribution to overall healthcare business success. There is also limited data regarding the combined acceptance and adoptions of these ICT tools by physicians, nurses and patients. When analyzing these assessments with a business-value approach, weaknesses assume a major role. Therefore, the next section attempts to minimize these gaps.

CONCEPTUAL FRAMEWORK AND HIPOTHESES

In the last three decades information technology in health care has existed and has gained widespread usage (Ammenwerth et al., 2004). Nevertheless, there are also hazards related with information technology in health care driven from the challenges and innovations adoption process and measurement.

For more than a decade it is widely believed that broad adoption of electronic medical record (EMR) systems will lead to major health care savings, reduce medical errors, and improve health care services (Marques, Oliveira, Dias, & Martins, 2011). However, the measurement of ICT impacts covers the assessment of the quality, value and their applications in the health care environment, the improvement of health information applications and the enabling of an evidence-based health informatics practice.

Thus, as it happens in other industries, the cost-benefits evaluation of IT implementation cannot be performed only base in traditional financial and accounting methods. Literature review revealed some intangible, indirect and even strategic benefits related to IT implementation that need to be considered, most of them common in IT domain (Grembergen & Amelinckx, 2002). Thus, there is undoubtedly a need to develop a better comprehension of ICT healthcare adoption impacts on global results.

The application of IT literature in this subject may shed some light on the measurement of the success of these initiatives (Romano Jr & Fjermestad, 2001). Yoon et al. (1995) suggested a framework based on five levels: business profitability, improved decision performance, level of system usage, perceived benefits and user satisfaction. The work of Eder and Igbaria (2001) analyses the impact of ICT on business performance and stressed that there are some indirect effects which need to be consider. More

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

recently, Bayo-Moriones and Lera-López (Bayo-Moriones et al., 2013; 2007) suggested that the positive impacts identified are not limited to increased productivity and reduced costs, influencing intermediate performance measures such as process efficiency, quality of services, organizational improvements, knowledge flows and customer satisfaction.

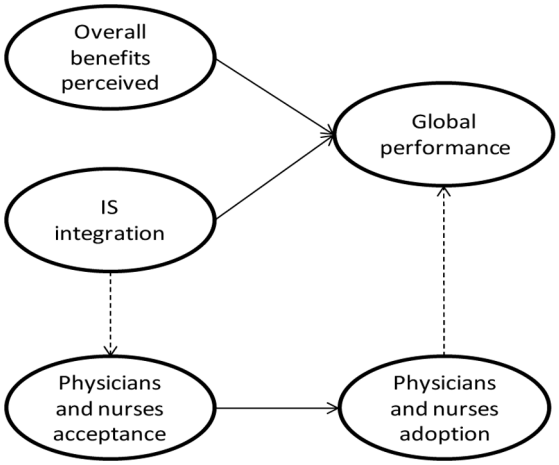
These last elements (quality of services, knowledge flows and customer satisfaction) can be improved by technology, but as mentioned previously, are dependent on end-user acceptance and adoption (I. Chang, 2007; Holden & Karsh, 2010; Jennings et al., 2015). From previous works, three sets of players need to be considered when adopting an HIS: patients (Hall & Dornan, 1988; Norum, Grev, Moen, Balteskard, & Holthe, 2003; Scalone et al., 2008); nurses (Abbott & Coenen, 2008; Verdegem & Verhoest, 2009) and physicians (Egea et al., 2015; Hunt, Haynes, Hanna, & Smith, 1998; Van Der Meijden, Tange, Troost, & Hasman, 2003; Verdegem & Verhoest, 2009).

In order to improve the success of information technology implementations, different models grounded in behavioral science or using theories that identify conditions and determinants of success are used (M.-Y. Chang et al., 2015; Hadgkiss & Renzaho, 2014; Heathfield et al., 1998; Holden & Karsh, 2010; Hunt et al., 1998; Mathieson, 1991; Norum et al., 2003). Nevertheless, as IT literature has evolved, these models have been tested separately and in a more-or-less restricted set of IT implementation procedures. In the healthcare setting, user behavior is influenced not only by the end-user personal framework, but also by a multifactor structure at the group, organizational and patient level. In this context, the defining of suitable multi-level intervention has been overlooked and this is the gap that the current research attempts to fulfill.

For this purpose, a significantly modified version of the original TAM was used, integrating the concepts presented by King & He (2006), and considering the relationships among the three main TAM variables—perceived ease of use, perceived usefulness and behavioral intention to use—and considering the three sets of players.

The research model, as shown in Figure 3, illustrates an examination of the productivity flows considering not only the operational process but also intangibles domains, which affects healthcare systems adoption and results. This framework also identifies metrics that can be used to measure the intangibles benefits. Each perspective is evaluated according to the appropriate metrics.

Figure 3. Conceptual Model



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The primary research question to consider is this: What are the benefits engages in performance, considering it as global productivity increasing, trough the implementation of ICT? And what is the impact of operational process (integrated IS) on global productivity performance? Besides measuring the convergence of a technological approach with a business value approach, our aim is to discover the direct nature of the impact of intangible benefits in healthcare firms' outcome.

The model is validated by using a structure equation model to test the hypotheses. The measurement model is done by using Confirmatory Factor Analysis (CFA). We will see later that the observable variables (indicators) we selected are measures of latent variables (factors). We assume that IT process and benefits driven from organization acceptance of IT have a direct effect on successful IS application. Therefore, we assume that the healthcare productivity performance gains are explained not only by the standard use of information systems, but also by the perceptions of organizational, technical and productivity gains.

Two sets of postulates were put together: first, considering as a theoretical base for the development of these postulates the literature on information systems and healthcare fields that evaluates the factors impacting healthcare performance; and, second, considering as a theoretical background the models related to technology adoption and acceptance.

The first is subdivided into three subsets and measures the impact of the three different kinds of benefits; following the results found by other researchers and bringing to HIS the measurement of intangible benefits already consider in other IT domains. The induce benefits are consider as all the benefits arise in terms of security, efficiency, reliability and staff satisfaction driven from the adoption of ICT. The financial benefits reflect the financial gains resulted from ICT implementation. The structural benefits consider the gains obtained from process modifications and operational cost reductions.

The second hypothesis seeks to confirm the inference that there is a positive impact of overall benefits on global healthcare productivity and consequently on healthcare firms' performance. The third hypothesis sought to identify a positive impact of the interoperability and integration of IS on global productivity performance, as an agent driving the optimize adoption and use of ICTs.

Another set of three questions was defined for each different set of players, based on the conceptual model present by King & He (2006). Therefore, the fourth hypothesis considers that the perceived ease of use of the implemented system impacts positively the acceptance of global HIS. The fifth proposition considers that healthcare providers' perception of the ICT usefulness impacts positively the acceptance of HIS. The last hypothesis advocates that behavioral intention, driven from personal past experiences and expectations, impacts positively HIS acceptance.

DATA TREATMENT AND RESULTS

To validate the first set of assumptions, data was collected from a sample of 834 European healthcare firms in a total of 17 countries (Belgium, Bulgaria, Czech Republic, Denmark, Finland, Greece, Spain, France, Hungary, Ireland, Italy, Latvia, The Netherlands, Norway, Poland, Sweden and UK) and covering seven dimensions of analysis: (1) ICT infrastructure and e-business software systems; (2) automated data exchange; (3) e-standards and interoperability issues; (4) Innovation activity of the company; (5) ICT skills requirements and ICT costs; (6) ICT impacts, drivers and inhibitors; and (7) background information about the company.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Initially we used a set of descriptive statistics that allowed for a greater sensitivity to the data (Sekaran, 2006). Such descriptive statistics also acted as guide for the multivariate statistics later on performed (Hair, Anderson, Tatham, & Black, 1998). Some notes about the sample used.

Distribution of firm size, measured by the number of employees, shows that almost half of the firms are micro- and small-size firms (around 50.7%). Around 99% of the firms in the sample have access to Internet and from those 28.7% use a wireless access. The importance of having ICT department at firm structure was measure and 49% of the sample currently employs ICT practitioners. The importance give to the lack of ICT department was measure by their impact on global ICT strategy and management (37.5% considers it a great deal), in ICT network architecture (56.4% considers it a huge concern), in ICT security (52.1% consider important) and maintenance (35.4% consider it important).

When asked about the IT solutions adopted by their firms, the highest percentage of respondents refers the use of pharmacy management system, as we can observe in the following figure 4.

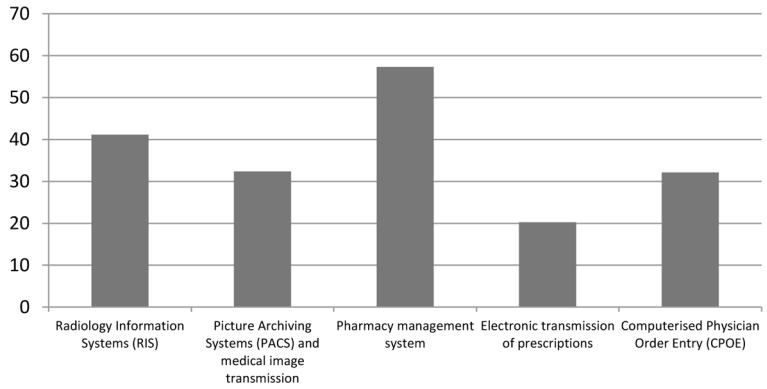
The model was estimated by the Maximum Likelihood method in the AMOS package. The model goodness of fit may be considered acceptable according to the values of some goodness-of-fit index, although the chi-square test statistic ($\chi 2 = 626{,}4$; df =321; p-value = 0,000) is significant, implying a bad fit. However, as is well known, this test has serious limitations — namely its dependence on the sample size and on the number of indicators. In general, for large sample sizes the chi-square statistic is significant, and in the present case the sample size is quite large (n = 834) (Hair et al., 1998). For that reason, it is usual to evaluate the goodness of the fit by a set of index, also presented in Figure 4.

After global model fit has been assessed, the numerical results were evaluated in order to test their support of the research question. The numerical results can be obtained directly from the path coefficients of the structural model presented in Figure 5. We refer to standardised coefficients that account for scale effects and serve as indicators of the relative importance of the variables.

The measures for global model fit included in Figure 5 suggest that our model fits the underlying data well (Hair et al., 1998) and all the paths were statistically significant.

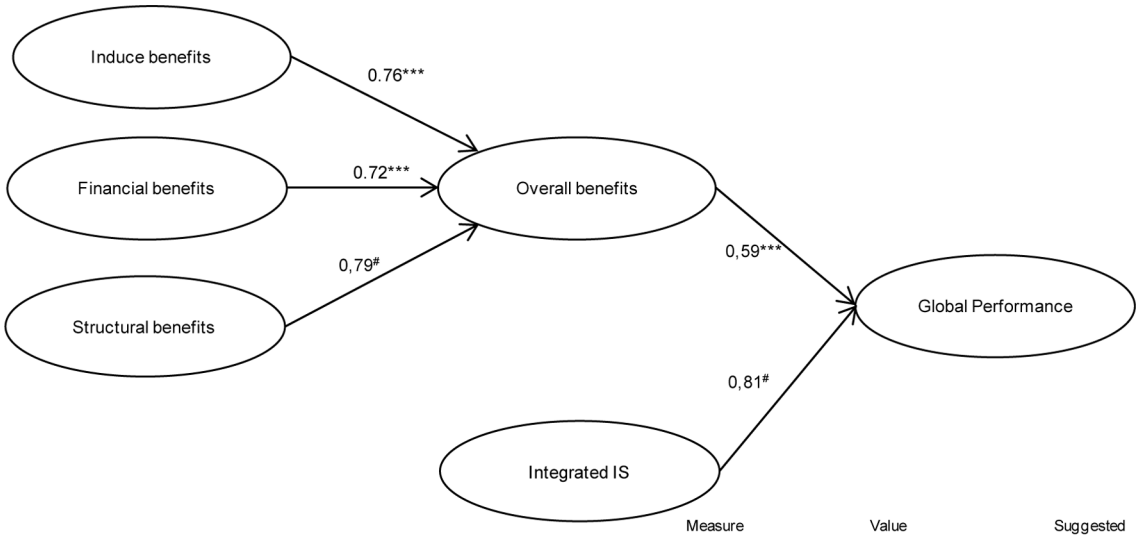
Our findings generally support the conceptual framework previously presented. The three dimensions used to compose the overall benefits achieved trough ICT adoption are all significant and validate trough the confirmatory factor analysis performed. Induce benefits, financial benefits and structural benefits explained 76 percent, 72 percent and 79 percent respectively regarding the variance in the overall benefits construct. As a result, hypothesis H1 is not rejected. A reference needs to be made regarding the relative similar value achieved in terms of financial benefits, reinforcing the perspective token in this research.

Figure 4. Percentage of adoption of different HIS



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Structural equation model and estimation results (model 1)



The results also allow concluding that the overall benefits contribute with 59 percent to global performance, so hypothesis H2 is not rejected.

The indirect effects existent in the model were also assessed. From those arises the strong relationship between interoperability of HIS and reduction of cost.

On the other hand, the results show that the construct "integrated IS", which is related to the interoperability between HIS present in the firm, contributes with 81 percent to global performance, implying that our hypothesis H3 is confirmed. A last note needs to be made relative to the value achieved in this path, which is greater than the relationship found between overall benefits and global productivity.

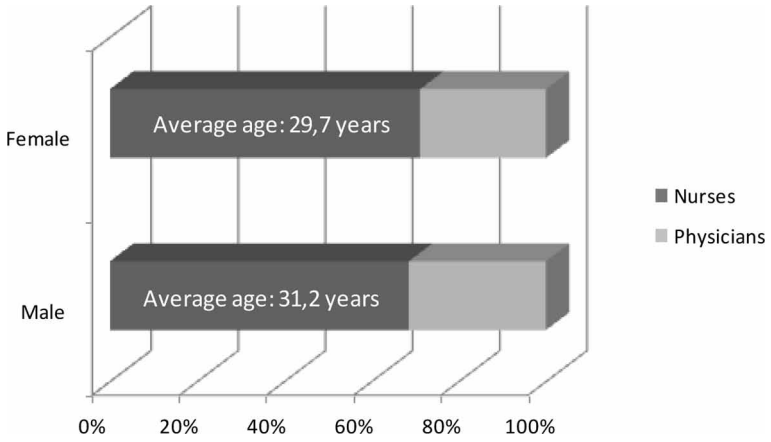
To validate the second subset of assumptions, data was collected from a sample of Portuguese physicians and nurses, covering four dimensions of analysis: 1) perceptions regarding the ease of use of current systems, 2) perceptions regarding impact of use, 3) behavioural intention to use ICT tools, and 4) past experience and demographics. Taking into consideration that the TAM will produce different results if the users being questioned have a) used the technology being tested previously and b) had a choice in whether to use the technology, the questionnaire will target healthcare providers with experience in the use of HIS.

Again, we used a set of descriptive statistics that allowed for a greater sensitivity to the data. As shown below, the sample has a higher representation of nurses, and in both profiles, the average age is quite low. This reflects the higher propensity of the millennial generation to use ICT tools, since the sample only gathered the data of individuals who actually used technology (see Figure 6).

All physicians use a computer at the hospital or clinic where they operate, other than the EMR, but only 23% report using it in private clinic. Looking closely at the perceived ease of use of electronic patient records, female physicians present a more positive perception of this tool: 50% consider EPR

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

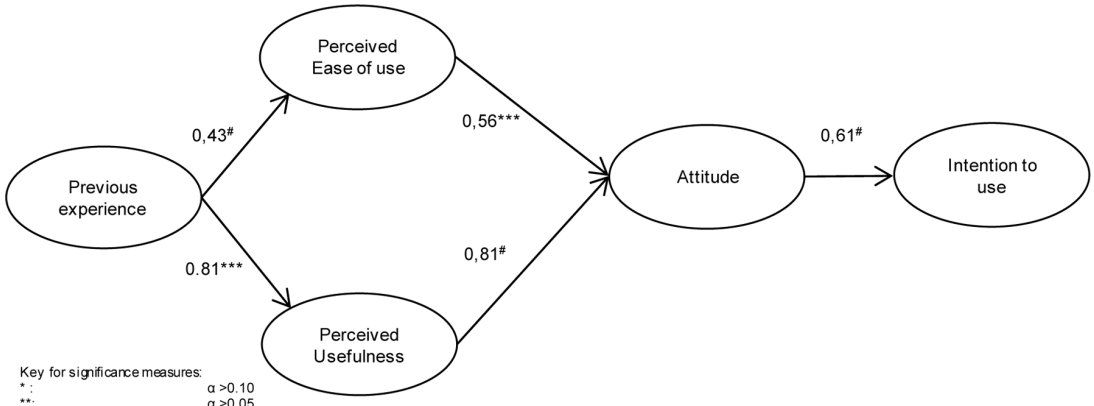
Figure 6. Demographic profile of healthcare providers' using ICT tools



very easy to use. Since there seems to exist a significant difference between female and male physicians (Gefen & Straub, 1997; Lu, Yu, Liu, & Yao, 2003; Venkatesh & Morris, 2000), the analysis was performed based on perceived gender differences. All nurses refer to using a computer for their daily records and show evidence of having a longer experience in doing so (average time of usage: 5 years). Regarding the physician-patient relationship, physicians are not convinced that the use of EMR improves the quality of medical care received by patients (44.7%), but 65.8% believe that it helps them perform better. Looking at nurses' answers, 56% believe that EMR does not improve the quality of medical care received by patients.

A second structural model was estimated by the Maximum Likelihood method in the AMOS package. After global model fit was assessed, the numerical results were evaluated in order to test their support of the research question. The numerical results can be obtained directly from the path coefficients of the structural model presented in Figure 7. We refer to standardised coefficients that account for scale effects and serve as indicators of the relative importance of the variables.

Figure 7. Structural equation model and estimation results (model 2)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

When performing multi-group analysis, the sample size of the groups being compared should be roughly equal. Since, in our sample, the number of nurses was larger, it wasn't possible to perform a multi-group analysis.

Once more, our findings, in general, support the conceptual framework. The dimensions used to compose the perceived ease of use and perceived usefulness are all significant and validate with the confirmatory factor analysis performed. Previous experience with the use of IT systems explained 43 and 81 percent of perceived ease of use and perceived usefulness, respectively. These results unveil some difficulties with the adoption of the technology and found that support of the attitude toward using ICT tools was positively correlated with intention behaviour. A note needs to be made relative to the value achieved in perceived usefulness and attitude towards use, which is greater than the relationship found between perceived ease of use and attitude.

DISCUSSION AND CONCLUSION

The digital era bring with it enormous challenges that healthcare firms can embrace, especially if we consider the facilities associated with the gathering of information about customers, suppliers, markets, and supply — and the easy processing of information regarding processes, and services, which can also be easily spread to the clients and other publics.

Even though academic researchers and practitioners alike praise information and communication technologies adoption (M.-Y. Chang et al., 2015; Hadgkiss & Renzaho, 2014; Marques et al., 2011), perhaps the most significant gap in the literature is the lack of large-scale empirical evidence showing that ICT makes a difference to the overall performance of the firm and determining which factors are critical to leverage overall performance.

The goal of the current study was twofold: 1) to identify the nature of the relationship between overall performance and these benefits, as well as the relationship with interoperability of HIS; and 2) to determine if the individual's acceptance of a technology is determined by its perceived usefulness and perceived ease of use. For this purpose, a set of propositions were developed: the first relies on exploring the nature of the benefits that can be achieved during the implementation and use of ICT tools; the second is based on the relationship between these benefits and healthcare performance; the third relies on the fact that most healthcare firms use more than one HIS, and for that reason, interoperability between all the systems can be crucial to global performance; the fourth and fifth predict that ease of use and usefulness will influence an individual's attitudes toward, intention to use, and acceptance of the technology. Therefore, our aim is to develop a new interpretation that offers fresh insights and understanding.

Given the nature of relationship between variables that is consider, the use of structural equation model is suitable. Based on the results of our two models, we can draw the following conclusions:

- Bayo-Moriones and Lera-López (2007) reported that the positive impacts occurred in an organization were not limited to increased productivity and reduced costs, but also influencing intermediate performance measures. The present work reinforces these ideas in the healthcare sector, presenting induces benefits achieved from ICT adoption.
- When comparing our results with the ones found by Egea et al. (2015), who also study European healthcare providers relationship with technologies, a positive trend seems to appear. This might be a sign that Europeans are more willing to adopt technologies which is fairly positive if con-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

sidering the downturn economic context and that those TI-solutions are quite effective in lower-resource settings.

- Our results confirmed some prior research regarding the relationship between interoperability of HIS and global performance (Macedo & Isaías, 2009; M. B. Tiago & Tiago, 2013), since we found a strong relationship between these constructs. In addition, indirect effects were also found, especially between interoperability of HIS and reduction of cost, which is an important dimension of global performance.
- Even though prior research on HIS field does not consider the intangible benefits of an adoption of ICT, we can empirically show that induce benefits influences the composition of total benefits. It means that this tendency, already observed in other IS domains, needs to be considered when evaluating the outcome of ICT investments.
- We validate the construct of global healthcare performance integrating measurement metrics besides the financial and accounting, following the path introduce by the work of Grembergen and Amelinckx (2002) and already tested in other domains (Maria Teresa Borges Tiago, Couto, Tiago, & Vieira, 2007; M.T.B. Tiago, Couto, Tiago, & Vieira, 2008).
- Our results also point to the existence of relevant benefits driven from structural processes which suffer improvements from the adoption of ICT, in line with Eder and Igbaria (2001), that had suggested the impact of some indirect effects on firms performance due to the use of ICT.
- The most surprising fact is, however, related to the relative contribution of financial benefits to total benefits. Since this is the most common dimension of analysis adopted by firms, we were expecting to find here the biggest contribution of all three dimensions. But, regardless your expectation, the impact found was strong but the smallest of all three.
- When looking at the results of the second sample, gender and age differences were found in terms of ICT adoption and benefit perceptions, which was already suggested in previous works (Gefen & Straub, 1997; Lu et al., 2003; Venkatesh & Morris, 2000).
- In the work of McGowan, et al (2012) focusing on social media usage by physicians, the authors stated that the frequency of usage is influenced primarily by positive attitudes toward the technology, perceiving that the technology is easy to use, and perceiving the technology to be useful to achieve better performance outcomes. Similar results were found in our sample.
- Our results also point to a lower relationship between perceived ease of use and attitude. These results might reflect the minimal training and customization of the IT systems used in the healthcare facilities in the sample.
- Our results didn't allow us to conclude if there was statistical evidence that physicians' and nurses have different behavior intentions regarding ICT. However, looking at the data we can infer that regarding to ICT acceptance and adoption the outcomes are not significantly different, which is quite unique when compare with the results from other studies regarding cost and quality of service provided by nurses and physicians (Jennings et al., 2015).

We caution readers to consider that the induce benefits and structural dimensions of HIS await future research and that the conclusions presented are related to the present samples used. For these reasons they cannot be generalized.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Issues, Controversies, Problems

It is often stated that healthcare is lagging in the rear of others sectors of our society when it comes to the adoption and use of ICT. If we look carefully at healthcare organizations strategic choices in terms of investment in this area, we still notice that the investments are consider as pilot projects and are disperse and with reduce or none IS architecture. This limited investment in ICT has been a rational choice since the measurement of the gains is still scarce. Besides, compensation mechanisms, public recognition and pressure, and consumer demand have not created a strong enough incentive for healthcare organizations to accept the financial costs and face up to the non-financial obstacles to ICT implementation.

However, the fast growth of web-based platforms that facilitate online social behavior has significantly modified the nature of human activities, habitats, and interactions (T. Tiago & Veríssimo, 2014). And service providers that neglect ICT investments are clearly loosing focus on consumer, regardless of the industry.

One of the major challenges concerning the ICT adoption is end users' acceptance. But who are these end users? Until a few years, the healthcare sector was focused on internal stakeholders. With the evolution of both medical and IT fields, the focus shift and started to integrate external stakeholders too.

From an internal perspective, if it is a common domain to recognize related to ICT adoption the reduction of medical errors, delays and misdiagnosis. But, it's also common to find references to the reduction of human relation or contact between physicians, patients and nurses. Thus, if for some this is a benefic consequence of ICT, for others, it is a lack of conformity that is damaging the outcome of the service. In short, if healthcare providers were rewarded for safety and quality results, they would be more likely to invest in ICT and promote an IT culture inside healthcare services stakeholders.

From a patient perspective, there is a wide range of benefits that can be achieved from the use of ICT, such as efficiency, convenience, richer and participate information, and social network application. As previously mentioned the Healthy 2.0 model answers to some of the challenges that the healthcare sector faces (Fabricio F. Costa, 2013). Again, this model relies on the end-user's acceptance and adoption, meaning that healthcare providers need to be motivated to adopt a digitally active posture.

Considering that for patients and relatives, through information communication technologies have a number of benefits including efficiency, convenience, richer and participative information, and accessibility, is relatively common to find these participants on social networks and in search motors looking for answers. With semantic web, big data and new OMICS technologies have been targeted, since it can enable a more efficient, low-cost and large-scale generation and analysis of biomedical data and personalize IT-driven healthcare service.

Especially considering that in the near future information sharing among the various health-specific social networks, and with the generic social networks (Facebook, Twitter), using Web 3.0 technology will improve the patient/provider relationship in all 4P's, paving the way for the development of new IT tools, such as Healthy 3.0.

However, the adoption of a Healthy 3.0 scenario requires that healthcare providers consider that the information and the digital sharing are trustworthy and valid. Additionally, it requires that they are ready to adopt these technologies and ultimately embrace them as a format for providing new services and lifelong learning.

In the meantime, governments and national service entities try to develop mechanisms through policy and legislation to minimize the effects of these inhibitors' and promote ICT standardization and capable pathways within the national service frameworks.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In essence, the ICT adoption process reveals six dimensions that need to be consider in order achieving success and points out to the importance of three major players in the process: physicians, patients or clients and nurses. If all these stakeholders are awareness of the benefits of ICT to the entire organization, the acceptance of ICT will be easier and the financial and productivity gains are greater. Thus, we can consider that the success of these initiatives relies on developing an ICT environment inside and outside the healthcare firm.

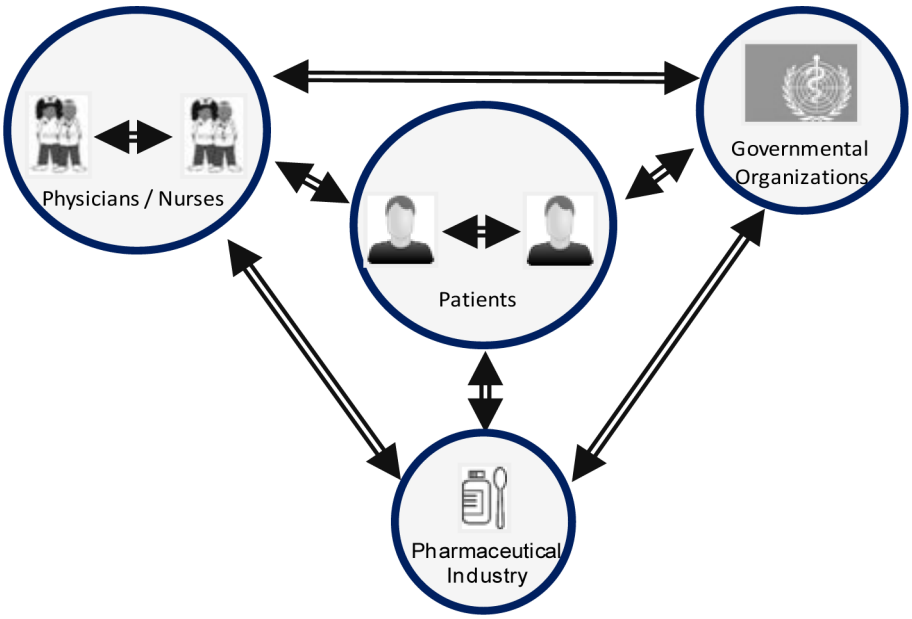
Solutions and Recommendations

Our comments are given with Portugal as main background for the healthcare sector together with a global understanding of the healthcare systems in other European countries and the USA. So, looking at the model present in this study several considerations can be made.

The first suggestion we make is related with acknowledge by all stakeholders the total benefits and not only the financial benefits of ICT implementation. In order to overall acknowledge of induce and structure benefits, a communication strategy needs to be set considering the three different major players: clients, physicians and nurses and their contact points with ICT. For instance, the communication with the last two publics can be executed trough training and public recognition of the outcomes. The success of this process relies on multiple actors' structure, as present below (see Figure 8):

Therefore, our suggestions included reinforcing harmonization activities between the existing entities of a national health system (NHS), starting with the clear identification of the players and their specific needs, followed by the development of specific ITS base on the identified needs. Afterwards, a certification procedure for e-health standards and adoptions should be promote, considering that such a process would allow interoperability of systems and the increasing of end-user orientation.

Figure 8. Healthy 3.0: healthcare digital dimensions



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Since the orientation should be focus on the end-user several actions should take place. Regarding health services clients, the emphasis needs to be on demystifying the concerns related to privacy and security. In this process, communication plays a crucial role and needs to reflect the organizations' commitment in leverage digital technologies to provide better, ubiquitous and less expensive services to customers.

The other two types of end-users should sense the support of their board, which is clear when auscultation during ITS development and implementation occurs; training and customization of ICT tools is provide; and global and personal efforts are recognize and compensate.

With the adoption of this matrix the intangible benefits can be identify and enlarge and interoperability of systems is assured. Nonetheless, to promote an easier acceptance process by inside players, the start timing and executive commitment is critical. This should be combined with the use of a team of expert innovators that will have as goals educate and promote the development of an IT culture.

FUTURE RESEARCH DIRECTIONS

This research produces some useful insights, leaving still a number of issues for future research. One of these issues is related to the possibility of extending the study in order to consider an inclusion of some dimensions of the technological acceptance model, such as technological personal readiness. Similarly, this study could be expanded covering the views and expectation of the three main players consider in this work: physicians, nurses and clients. The analysis of their perspective on this matter can reinforce some of the conclusions achieved, challenge some of the outcomes of this work and certainly contribute to a deeper understanding of healthcare IS evolution and prospective.

Another path to follow regards the new ways of healthcare delivery, the new public and their solicitations. This will require additional IS functionalities, been an interesting field for future research.

CONCLUSION

The results generated some interesting findings. First, the data support our conceptualization for the global benefits construct. Within global benefits elements, all have a positive impact on the maximization of firms ICT-driven benefits and consequently global performance. Secondly, the findings allow us to conclude that induce benefits have a positive indirect impact on global productivity. Thus, financial and intangible benefits appear to greater than what was usual accounted, paying back the initial investments made by the healthcare firms. Nonetheless, the size of the effect for this relationship was smaller than would be expected.

According to these results, the concept of integrated systems as an important dimension of overall results equation is reinforced. So, managers should consider the use of HIS competencies to improve everyday services process. Namely, taking advantage of the benefits such as reducing structural and operational costs, increasing employees' motivation and satisfaction, improving services quality as pointed out by others researches.

More, healthcare providers should be motivated to adopt ICT tools as an important opportunity to engage customers and perform a high-quality service.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Abbott, P. A., & Coenen, A. (2008). Globalization and advances in information and communication technologies: The impact on nursing and health. *Nursing Outlook, 56*(5), 238-246. e232.

Ahmad, F., Hudak, P. L., Bercovitz, K., Hollenberg, E., & Levinson, W. (2006). Are Physicians Ready for Patients With Internet-Based Health Information? *Journal of Medical Internet Research*, *8*(3), e22. doi:10.2196/jmir.8.3.e22 PMID:17032638

Ammenwerth, E., Brender, J., Nykanen, P., Prokosch, H. U., Rigby, M., & Talmon, J. (2004). Visions and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. *International Journal of Medical Informatics*, *73*(6), 479–491. doi:10.1016/j.ijmedinf.2004.04.004 PMID:15171977

Ammenwerth, E., Gräber, S., Herrmann, G., Bürkle, T., & König, J. (2003). Evaluation of health information systems--problems and challenges. *International Journal of Medical Informatics*, *71*(2-3), 125–135. doi:10.1016/S1386-5056(03)00131-X PMID:14519405

Ammenwerth, E., Schnell-Inderst, P., Machan, C., & Siebert, U. (2008). The effect of electronic prescribing on medication errors and adverse drug events: A systematic review. *Journal of the American Medical Informatics Association*, *15*(5), 585–600. doi:10.1197/jamia.M2667 PMID:18579832

Anderson, E. W., Fornell, C., & Rust, R. T. (1997). Customer satisfaction, productivity, and profitability: Differences between goods and services. *Marketing Science*, *16*(2), 129–145. doi:10.1287/mksc.16.2.129

Bates, D. W., Leape, L. L., Cullen, D. J., Laird, N., Petersen, L. A., Teich, J. M., & ... (1998). Effect of Computerized Physician Order Entry and a Team Intervention on Prevention of Serious Medication Errors. *Journal of the American Medical Association*, *280*(15), 1311–1316. doi:10.1001/jama.280.15.1311 PMID:9794308

Bayo-Moriones, A., Billón, M., & Lera-López, F. (2013). Perceived performance effects of ICT in manufacturing SMEs. *Industrial Management & Data Systems*, *113*(1), 117–135. doi:10.1108/02635571311289700

Bayo-Moriones, A., & Lera-López, F. (2007). A firm-level analysis of determinants of ICT adoption in Spain. *Technovation*, *27*(6-7), 352–366. doi:10.1016/j.technovation.2007.01.003

Brannigan, V. (1985). Acceptance Testing-The Critical Problem in Software Acquisition. *IEEE Transactions on Bio-Medical Engineering*, *BME-32*(4), 295–299. doi:10.1109/TBME.1985.325451 PMID:3838735

Brothers, K. B., & Rothstein, M. A. (2015). Ethical, legal and social implications of incorporating personalized medicine into healthcare. *Personalized Medicine*, *12*(1), 43–51. doi:10.2217/pme.14.65 PMID:25601880

Budden, C. B., Anthony, J. F., Budden, M. C., & Jones, M. A. (2011). Managing the evolution of a revolution: Marketing implications of internet media usage among college students. *College Teaching Methods & Styles Journal (CTMS), 3*(3).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Buntin, M. B., Burke, M. F., Hoaglin, M. C., & Blumenthal, D. (2011). The benefits of health information technology: A review of the recent literature shows predominantly positive results. *Health Affairs*, *30*(3), 464–471. doi:10.1377/hlthaff.2011.0178 PMID:21383365

Chang, I., Hwang, H.-G., Hung, M.-C., Lin, M.-H., & Yen, D. C. (2007). Factors affecting the adoption of electronic signature: Executives' perspective of hospital information department. *Decision Support Systems*, *44*(1), 350–359. doi:10.1016/j.dss.2007.04.006

Chang, M.-Y., Pang, C., Tarn, J. M., Liu, T.-S., & Yen, D. C. (2015). Exploring user acceptance of an e-hospital service: An empirical study in Taiwan. *Computer Standards & Interfaces*, *38*, 35–43. doi:10.1016/j.csi.2014.08.004

Classen, D. C., Resar, R., Griffin, F., Federico, F., Frankel, T., Kimmel, N., & ... (2011). 'Global trigger tool'shows that adverse events in hospitals may be ten times greater than previously measured. *Health Affairs*, *30*(4), 581–589. doi:10.1377/hlthaff.2011.0190 PMID:21471476

Costa, F. F. (2013). Social networks, web-based tools and diseases: Implications for biomedical research. *Drug Discovery Today*, *18*(5), 272–281. doi:10.1016/j.drudis.2012.10.006 PMID:23103282

Darkins, A., Foster, L., Anderson, C., Goldschmidt, L., & Selvin, G. (2013). The design, implementation, and operational management of a comprehensive quality management program to support national telehealth networks. *Telemedicine Journal and e-Health*, *19*(7), 557–564. doi:10.1089/tmj.2012.0263 PMID:23705961

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003. doi:10.1287/mnsc.35.8.982

Davis, T. C., Jackson, R. H., George, R. B., Long, S. W., Talley, D., Murphy, P. W., & ... (1993). Reading ability in patients in substance misuse treatment centers. *Substance Use & Misuse*, *28*(6), 571–582. doi:10.3109/10826089309039648 PMID:8098019

Dewett, T., & Jones, G. R. (2001). The role of information technology in the organization: A review, model, and assessment. *Journal of Management*, *27*(3), 313–346. doi:10.1177/014920630102700306

Dhillon, J. S., Wünsche, B. C., & Lutteroth, C. (2013). *An online social-networking enabled telehealth system for seniors: a case study*. Paper presented at the Fourteenth Australasian User Interface Conference.

Drapeau, M., & Wells, I. (2009). *Social software and national security: An initial net assessment*. DTIC Document.

Eder, L. B., & Igbaria, M. (2001). Determinants of intranet diffusion and infusion. *Omega*, *29*(3), 233–242. doi:10.1016/S0305-0483(00)00044-X

Egea, J. M. O., González, M. V. R., & Menéndez, M. R. (2015). *E-Health Services Adoption by European General Practitioners Assessing the Different Roles of Marketing Theory and Practice in the Jaws of Economic Uncertainty*. Springer.

Fischer, S., & Soyez, K. (2015). *Trick or Treat: Assessing Health 2.0 and Its Prospects for Patients, Providers and Society Challenges and Opportunities in Health Care Management*. Springer.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Flagle, C., Grémy, F., & Perry, S. (1991). *Assessment of Medical Informatics Technology*. Paper presented at the Joint Working Conference.

Fracolli, L. A., Gomes, M. F. P., Nabão, F. R. Z., Santos, M. S., Cappellini, V. K., & Almeida, A. C. C. (2014). Primary health care assessment tools: A literature review and metasynthesis. *Ciencia & Saude Coletiva*, *19*(12), 4851–4860. doi:10.1590/1413-812320141912.00572014 PMID:25388193

Gefen, D., & Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the technology acceptance model. *Management Information Systems Quarterly*, *21*(4), 389–400. doi:10.2307/249720

Grembergen, W., & Amelinckx, I. (2002). *Measuring and Managing E-business Projects through the Balanced Scorecard*. Paper presented at the 35th Hawaii International Conference on System Science, Hawaii, HI. 10.1109/HICSS.2002.994395

Gritzalis, D. (1997). A baseline security policy for distributed healthcare information systems. *Computers & Security*, *16*(8), 709–719. doi:10.1016/S0167-4048(97)00009-6

Hadgkiss, E. J., & Renzaho, A. M. (2014). The physical health status, service utilisation and barriers to accessing care for asylum seekers residing in the community: A systematic review of the literature. *Australian Health Review*, *38*(2), 142–159. PMID:24679338

Hair, J., Anderson, R., Tatham, R., & Black, W. (Eds.). (1998). *Multivariate Data Analysis*. Prentice Hall.

Hall, J. A., & Dornan, M. C. (1988). Meta-analysis of satisfaction with medical care: Description of research domain and analysis of overall satisfaction levels. *Social Science & Medicine*, *27*(6), 637–644. doi:10.1016/0277-9536(88)90012-3 PMID:3067359

Heathfield, H., Pitty, D., & Hanka, R. (1998). Evaluating information technology in health care: Barriers and challenges. *BMJ (Clinical Research Ed.)*, *316*(7149), 1959–1961. doi:10.1136/bmj.316.7149.1959 PMID:9641938

Henderson, C., Knapp, M., Fernández, J.-L., Beecham, J., Hirani, S., Beynon, M., . . ., Doll, H. (2012). Cost-effectiveness of telehealth in people with social care needs: the Whole Systems Demonstrator cluster randomised trial. *International Journal of Integrated Care, 12*(Suppl1).

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs*, *24*(5), 1103–1117. doi:10.1377/hlthaff.24.5.1103 PMID:16162551

Hippisley-Cox, J., Pringle, M., Cater, R., Wynn, A., Hammersley, V., Coupland, C., & ... (2003). The electronic patient record in primary care-regression or progression? A cross sectional study. *British Medical Journal*, *326*(7407).

Holden, R. J., & Karsh, B. T. (2010). The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics*, *43*(1), 159–172. doi:10.1016/j.jbi.2009.07.002 PMID:19615467

Hu, P. J., Chau, P. Y. K., Sheng, O. R. L., & Tam, K. Y. (1999). Examining the technology acceptance model using physician acceptance of telemedicine technology. *Journal of Management Information Systems*, *16*(2), 91–112. doi:10.1080/07421222.1999.11518247

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hung, S.-Y., Hung, W.-H., Tsai, C.-A., & Jiang, S.-C. (2010). Critical factors of hospital adoption on CRM system: Organizational and information system perspectives. *Decision Support Systems*, *48*(4), 592–603. doi:10.1016/j.dss.2009.11.009

Hunt, D. L., Haynes, R. B., Hanna, S. E., & Smith, K. (1998). Effects of computer-based clinical decision support systems on physician performance and patient outcomes. *Journal of the American Medical Association*, *280*(15), 1339. doi:10.1001/jama.280.15.1339 PMID:9794315

Inklaar, R., O'Mahony, M., & Timmer, M. (2005). ICT and Europe's productivity performance. industry level growth account comparisons with the United States. *Review of Income and Wealth*, *51*(4), 505–536. doi:10.1111/j.1475-4991.2005.00166.x

Jennings, N., Clifford, S., Fox, A. R., O'Connell, J., & Gardner, G. (2015). The impact of nurse practitioner services on cost, quality of care, satisfaction and waiting times in the emergency department: A systematic review. *International Journal of Nursing Studies*, *52*(1), 421–435. doi:10.1016/j.ijnurstu.2014.07.006 PMID:25443302

Jensen, J., Tøndel, I. A., Jaatun, M. G., Meland, P. H., & Andresen, H. (2009). *Reusable Security Requirements for Healthcare Applications*. Academic Press.

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, *43*(6), 740–755. doi:10.1016/j.im.2006.05.003

Kumar, R., Novak, J., & Tomkins, A. (2010). Structure and evolution of online social networks. *Link Mining: Models, Algorithms, and Applications*, 337-357.

Lai, L. S., & Turban, E. (2008). Groups formation and operations in the Web 2.0 environment and social networks. *Group Decision and Negotiation*, *17*(5), 387–402. doi:10.100710726-008-9113-2

Lapointe, L., & Rivard, S. (2006). Getting physicians to accept new information technology: Insights from case studies. *Canadian Medical Association Journal*, *174*(11), 1573–1578. doi:10.1503/cmaj.050281 PMID:16717265

Lau, A. S. (2011). Hospital-based nurses' perceptions of the adoption of Web 2.0 tools for knowledge sharing, learning, social interaction and the production of collective intelligence. *Journal of Medical Internet Research*, *13*(4), e92. doi:10.2196/jmir.1398 PMID:22079851

Lee, M. A., & Yom, Y. H. (2007). A comparative study of patients' and nurses' perceptions of the quality of nursing services, satisfaction and intent to revisit the hospital: A questionnaire survey. *International Journal of Nursing Studies*, *44*(4), 545–555. doi:10.1016/j.ijnurstu.2006.03.006 PMID:16687147

Liao, S. H. (2003). Knowledge management technologies and applications - literature review from 1995 to 2002. *Expert Systems with Applications*, *25*(2), 155–164. doi:10.1016/S0957-4174(03)00043-5

Liu, M.-S., & Liu, N.-C. (2008). Sources of knowledge acquisition and patterns of knowledge-sharing behaviors--An empirical study of Taiwanese high-tech firms. *International Journal of Information Management*, *28*(5), 423–432. doi:10.1016/j.ijinfomgt.2008.01.005

Lu, J., Yu, C. S., Liu, C., & Yao, J. E. (2003). Technology acceptance model for wireless Internet. *Internet Research*, *13*(3), 206–222. doi:10.1108/10662240310478222

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lucas, H. (2008). Information and communications technology for future health systems in developing countries. *Social Science & Medicine*, *66*(10), 2122–2132. doi:10.1016/j.socscimed.2008.01.033 PMID:18343005

Macedo, M., & Isaías, P. (2009). Patient Standardization Identification as a Healthcare Issue. *Human Interface and the Management of Information. Information and Interaction*, 566-575.

Marques, A., Oliveira, T., Dias, S. S., & Martins, M. F. O. (2011). Medical Records System Adoption in European Hospitals. *The Electronic Journal Information Systems Evaluation*, *14*(1), 89–99.

Mathieson, K. (1991). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research*, *2*(3), 173–191. doi:10.1287/isre.2.3.173

McGowan, B. S., Wasko, M., Vartabedian, B. S., Miller, R. S., Freiherr, D. D., & Abdolrasulnia, M. (2012). Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. *Journal of Medical Internet Research*, *14*(5), e117. doi:10.2196/jmir.2138 PMID:23006336

McKee, D. O., Varadarajan, P. R., & Pride, W. M. (1989). Strategic Adaptability and Firm Performance: A Market-Contingent Perspective. *Journal of Marketing*, *53*(3), 21–35. doi:10.2307/1251340

Murphy, G. D. (2010). *Using Web 2.0 tools to facilitate knowledge transfer in complex organisational environments: a primer*. Academic Press.

Norum, J., Grev, A., Moen, M. A., Balteskard, L., & Holthe, K. (2003). Information and communication technology (ICT) in oncology. Patients' and relatives' experiences and suggestions. *Supportive Care in Cancer*, *11*(5), 286–293. PMID:12690539

Papastathopoulou, P., Avlonitis, G. J., & Panagopoulos, N. G. (2007). Intraorganizational information and communication technology diffusion: Implications for industrial sellers and buyers. *Industrial Marketing Management*, *36*(3), 322–336. doi:10.1016/j.indmarman.2005.10.002

Peansupap, V., & Walker, D. (2005). Factors enabling information and communication technology diffusion and actual implementation in construction organisations. *Electronic Journal of Information Technology in Construction*, *10*, 193–218.

Pilat, D. (2004). The ICT productivity paradox: Insights from micro data. *OECD Economic Studies*, *38*(1), 37–65.

Pirnejad, H., Bal, R., Stoop, A. P., & Berg, M. (2007). Inter-organisational communication networks in healthcare: Centralised versus decentralised approaches. *International Journal of Integrated Care*, *7*(2). PMID:17627296

Rahimi, B., & Vimarlund, V. (2007). Methods to evaluate health information systems in healthcare settings: A literature review. *Journal of Medical Systems*, *31*(5), 397–432. doi:10.100710916-007-9082-z PMID:17918694

Raitoharju, R. (2005). *When Acceptance is not Enough-Taking TAM-Model into Healthcare*. Academic Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Richter, A., Stocker, A., Müller, S., & Avram, G. (2013). Knowledge management goals revisited: A cross-sectional analysis of social software adoption in corporate environments. *Vine*, *43*(2), 132–148. doi:10.1108/03055721311329927

Romano, N. C. Jr, & Fjermestad, J. (2001). Electronic commerce customer relationship management: An assessment of research. *International Journal of Electronic Commerce*, *6*(2), 61–113.

Ruland, C. M., & Ravn, I. H. (2003). Usefulness and effects on costs and staff management of a nursing resource management information system. *Journal of Nursing Management*, *11*(3), 208–215. doi:10.1046/j.1365-2834.2003.00381.x PMID:12694368

Scalone, L., Mantovani, L. G., Krol, M., Rofail, D., Ravera, S., Grazia Bisconte, M., & ... (2008). Costs, quality of life, treatment satisfaction and compliance in patients with -thalassemia major undergoing iron chelation therapy: The ITHACA study. *Current Medical Research and Opinion*, *24*(7), 1905–1917. PMID:18507891

Schillewaert, N., Ahearne, M. J., Frambach, R. T., & Moenaert, R. K. (2005). The adoption of information technology in the sales force. *Industrial Marketing Management*, *34*(4), 323–336. doi:10.1016/j.indmarman.2004.09.013

Sekaran, U. (2006). *Research methods for business*. Wiley India Pvt. Ltd.

Sharma, S., Kilian, R., & Leung, F.-H. (2014). Health 2.0—Lessons Learned Social Networking With Patients for Health Promotion. *Journal of Primary Care & Community Health*.

Sheng, M., Chang, S.-Y., Teo, T. S., & Lin, Y.-F. (2013). Knowledge Barriers, Knowledge Transfer, and Innovation Competitive Advantage in Healthcare Settings. *Management Decision*, *51*(3), 1–1. doi:10.1108/00251741311309607

Steventon, A., Bardsley, M., Billings, J., Dixon, J., Doll, H., Hirani, S., . . . Henderson, C. (2012). Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. *BMJ: British Medical Journal, 344*.

Tekli, J., Rjeily, A. A., Chbeir, R., Tekli, G., Houngue, P., Yetongnon, K., et al. (2013). *Semantic to intelligent web era: building blocks, applications, and current trends*. Paper presented at the Fifth International Conference on Management of Emergent Digital EcoSystems.

Tiago, M. B., & Tiago, F. (2013). *Roadmaps for Future ICT Use in the Health Sector Handbook of Research on ICTs and Management Systems for Improving Efficiency in Healthcare and Social Care*. IGI Global. doi:10.4018/978-1-4666-3990-4.ch035

Tiago, M. T. B., Couto, J. P., Tiago, F., & Vieira, A. C. (2007). Knowledge management:An overview of European reality. *Management Research News*, *30*(2), 100–114. doi:10.1108/01409170710722946

Tiago, M. T. B., Couto, J. P. A., Tiago, F. G., & Vieira, J. A. C. (2008). The missing link between virtual CRM and e-business performance1. *International Journal of Business and Management*, *2*(1).

Tiago, T., & Veríssimo, J. (2014). Digital marketing and social media: Why bother? *Business Horizons*, *57*(6), 703–708. doi:10.1016/j.bushor.2014.07.002

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation*. Lexington Books.

Van Ark, B., Inklaar, R., & McGuckin, R. H. (2003). ICT and productivity in Europe and the United States Where do the differences come from? *CESifo Economic Studies*, *49*(3), 295–318. doi:10.1093/cesifo/49.3.295

Van Der Meijden, M., Tange, H., Troost, J., & Hasman, A. (2003). Determinants of success of inpatient clinical information systems: A literature review. *Journal of the American Medical Informatics Association*, *10*(3), 235–243. doi:10.1197/jamia.M1094 PMID:12626373

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *Management Information Systems Quarterly*, *24*(1), 115–139. doi:10.2307/3250981

Verdegem, P., & Verhoest, P. (2009). Profiling the non-user: Rethinking policy initiatives stimulating ICT acceptance. *Telecommunications Policy*, *33*(10-11), 642–652. doi:10.1016/j.telpol.2009.08.009

Wang, S. J., Middleton, B., Prosser, L. A., Bardon, C. G., Spurr, C. D., Carchidi, P. J., & ... (2003). A cost-benefit analysis of electronic medical records in primary care. *The American Journal of Medicine*, *114*(5), 397–403. doi:10.1016/S0002-9343(03)00057-3 PMID:12714130

Ware, J. E., Davies, A. R., & Stewart, A. L. (1977). *The measurement and meaning of patient satisfaction: a review of the literature*. Rand Corporation.

Wild, R., & Griggs, K. (2008). A model of information technology opportunities for facilitating the practice of knowledge management. *Vine*, *38*(4), 490–506. doi:10.1108/03055720810917732

Yi, M. Y., & Hwang, Y. (2003). Predictingthe use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. Int. *J. Human-Computer Studies*, *59*(4), 431–449. doi:10.1016/S1071-5819(03)00114-9

Yoon, Y., Guimaraes, T., & O'Neal, Q. (1995). Exploring the factors associated with expert systems success. *Management Information Systems Quarterly*, *19*(1), 83–106. doi:10.2307/249712

Zhu, K., & Kraemer, K. (2002). E-commerce Metrics for Net-enhanced Organizations: Assessing the Value of e-commerce to Firm Performance in the Manufacturing Sector. *Information Systems Research*, *13*(3), 275–295. doi:10.1287/isre.13.3.275.82

Zhu, K., Kraemer, K., & Dedrick, J. (2004). Information Technology payoff in e-business environments: Na international perspective of value creation of e-business in the financial service industry. *Journal of Management Information Systems*, *21*(1), 17–54.

ADDITIONAL READING

Andersen, S. T., & Jansen, A. (2011). Innovation in ICT-Based Health Care Provision. [IJHISI]. *International Journal of Healthcare Information Systems and Informatics*, *6*(2), 14–27. doi:10.4018/jhisi.2011040102

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Anderson, J., & Aydin, C. (2005). Overview: Theoretical perspectives and methodologies for the evaluation of healthcare information systems. *Evaluating the Organizational Impact of Healthcare Information Systems*, 5-29.

Anderson, L. M., Scrimshaw, S. C., Fullilove, M. T., Fielding, J. E., & Normand, J. (2003). Culturally competent healthcare systems* 1: A systematic review. *American Journal of Preventive Medicine*, *24*(3), 68–79. doi:10.1016/S0749-3797(02)00657-8 PMID:12668199

Andreassen, H. K. (2010). What does an e-mail address add?-Doing health and technology at home. *Social Science & Medicine*. PMID:21208702

Antheunis, M. L., Tates, K., & Nieboer, T. E. (2013). Patients' and health professionals' use of social media in health care: Motives, barriers and expectations. *Patient Education and Counseling*, *92*(3), 426–431. doi:10.1016/j.pec.2013.06.020 PMID:23899831

Bakker, A. R. (2002). Health care and ICT, partnership is a must. *International Journal of Medical Informatics*, *66*(1-3), 51–57. doi:10.1016/S1386-5056(02)00036-9 PMID:12453558

Black, E., Light, J., Black, N. P., & Thompson, L. (2013). Online Social Network Use by Health Care Providers in a High Traffic Patient Care Environment. *Journal of Medical Internet Research*, *15*(5), e94. doi:10.2196/jmir.2421 PMID:23685530

Chaudhry, B., Wang, J., Wu, S., Maglione, M., Mojica, W., Roth, E., & ... (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine*, *144*(10), 742. doi:10.7326/0003-4819-144-10-200605160-00125 PMID:16702590

Cheong, L. H., Armour, C. L., & Bosnic-Anticevich, S. Z. (2013). Primary health care teams and the patient perspective: A social network analysis. *Research in Social & Administrative Pharmacy*, *9*(6), 741–757. doi:10.1016/j.sapharm.2012.12.003 PMID:23395532

Doarn, C. R., Yellowlees, P., Jeffries, D. A., Lordan, D., Davis, S., Hammack, G., & ... (2008). Societal drivers in the applications of telehealth. *Telemedicine Journal and e-Health*, *14*(9), 998–1002. doi:10.1089/tmj.2008.0111 PMID:19035816

Doherty, R. B. (2010). The certitudes and uncertainties of health care reform. *Annals of Internal Medicine*, *152*(10), 679. doi:10.7326/0003-4819-153-1-201007060-00235 PMID:20378676

Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P., Beyene, J., & ... (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes. *Journal of the American Medical Association*, *293*(10), 1223. doi:10.1001/jama.293.10.1223 PMID:15755945

Grajales, F. J. III, Sheps, S., Ho, K., Novak-Lauscher, H., & Eysenbach, G. (2014). Social Media: A Review and Tutorial of Applications in Medicine and Health Care. *Journal of Medical Internet Research*, *16*(2). PMID:24518354

Hamid, A., & Sarmad, A. (2011). Multi-Dimensional Criteria for the Evaluation of E-Health Services.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hansen, D. P., Gurney, P., Morgan, G., & Barraclough, B. (2011). The Australian e-Health Research Centre: Enabling the health care information and communication technology revolution. *The Medical Journal of Australia*, *194*(4), S5–S7. PMID:21401490

Hippisley-Cox, J., & Pringle, M. (2004). Prevalence, care, and outcomes for patients with diet-controlled diabetes in general practice: Cross sectional survey. *Lancet*, *364*(9432), 423–428. doi:10.1016/S0140-6736(04)16765-2 PMID:15288740

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ (Clinical Research Ed.)*, *330*(7494), 765. doi:10.1136/bmj.38398.500764.8F PMID:15767266

Mars, M., & Scott, R. E. (2010). Global e-health policy: A work in progress. *Health Affairs*, *29*(2), 237–243. doi:10.1377/hlthaff.2009.0945 PMID:20348067

Moffatt, J., Eley, D., & Eley, J. M. D. (2011). Barriers to the up-take of telemedicine in Australia–a view from providers. *Rural and Remote Health*, *11*(1581). PMID:21385004

Nickelson, D. W. (1998). Telehealth and the evolving health care system: Strategic opportunities for professional psychology. *Professional Psychology, Research and Practice*, *29*(6), 527–535. doi:10.1037/0735-7028.29.6.527

Orszag, P. R., & Emanuel, E. J. (2010). Health care reform and cost control. *The New England Journal of Medicine*, *363*(7), 601–603. doi:10.1056/NEJMp1006571 PMID:20554975

Parish, M. B., & Yellowlees, P. (2014). The Rise of Person-Centered Healthcare and the Influence of Health Informatics and Social Network Applications on Mental Health Care. In *Mental Health Informatics* (pp. 17–39). Springer Berlin Heidelberg. doi:10.1007/978-3-642-38550-6_2

Park, E., Baek, S., Ohm, J., & Chang, H. J. (2014). Determinants of player acceptance of mobile social network games: An application of extended technology acceptance model. *Telematics and Informatics*, *31*(1), 3–15. doi:10.1016/j.tele.2013.07.001

Sánchez, J. L., Savin, S., & Vasileva, V. (2005). *Key success factors in implementing electronic medical records in University Hospital of Rennes. L'Ecole Nationale de la Santé Publique* (pp. 1–59). Rennes, Rennes, France: National School of Public Health.

While, A., & Dewsbury, G. (2011). Nursing and information and communication technology (ICT): A discussion of trends and future directions. *International Journal of Nursing Studies*, *48*(10), 1302–1310. doi:10.1016/j.ijnurstu.2011.02.020 PMID:21474135

KEY TERMS AND DEFINITIONS

Healthy 3.0: Adoption of Web 3.0 technology-based tools to improve the patient/provider relationship.

HIS Measurement: Methods to evaluate outputs and outcomes of the use of IT-based systems in medical informatics.

ICT Acceptance: Individual and firms 'attitudes regarding technology and their intention to use.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ICT Induce Benefits: Intangible benefits achieved by the adoption and use of ICT in healthcare.

ICT Performance: ICT can give significant boost productivity gains of a firm, which can be measure through financial and non-financial metrics.

ICT ROI: Financial accountable benefits driven from the value gains and cost reduction.

ICT Support: Technical and training support in daily basis and board commitment to ICT use, measure by recognition of efforts and compensation.

This research was previously published in Reshaping Medical Practice and Care with Health Information Systems edited by Ashish Dwivedi; pages 287-322; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 61
Informed Consent in Healthcare:
A Study Case of Genetic Services

Bo Yu
George Mason University, USA

Duminda Wijesekera
George Mason University, USA

Paulo Costa
George Mason University, USA

ABSTRACT

Informed consents, either for treatment or sensitive information use/disclosure, that protect the privacy of patient/participant information subject to law that in certain circumstances may override patient wishes, are mandatory practice in healthcare. Similarly, for protecting and respecting research participants, informed consents are also prerequisite for human subjects research. Although the healthcare industry has widely adopted Electronic Medical Record (EMR) systems, consents are still obtained and stored primarily on paper or scanned electronic documents. Integrating a consent management system for different purposes into an EMR system involves various implementation challenges. A case study, informed consent for genetic services, is used to show how genetic informed consents placed new challenges on the traditional ethical standards of informed consent, and how appropriate consents can be electronically obtained and automatically enforced using a system that combines medical workflows and hierarchically, ontologically motivated rule enforcement. Finally, this chapter describes an implementation that uses the open-source software-based addition of these components to an open-source EMR system, so that existing systems do not need to be scrapped or otherwise rendered obsolete.

DOI: 10.4018/978-1-7998-1204-3.ch061

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The healthcare industry seeks quality improvement strategies to provide standardized, safer and more efficient patient services as well as medical research, focusing on patient safety, quality of care, risk management, data security and privacy preservation. The term "informed consent", first used by a California appeals court in 1957 (Cal. App. Ct. 1957), has become a vital ethical doctrine in medical practice. The American Medical Association states that "obtaining informed consent is an ethical obligation of the practice of medicine and a legal requirement per statute and case law in all 50 States" (O'Leary, 2010). In healthcare domain, informed consent has been considered one of the effective strategies to achieve the above goals. Herein the authors focus on three different categories: consent for medical treatment; consent for sensitive medical information use and disclosure; and consent for human subjects research.

In the last several decades, evolving computer-based Information Technology (IT) in healthcare has changed how computing systems support healthcare service delivery, with examples including Electronic Medical Record (EMR) systems and Personal Health Record (PHR) systems. The U.S government plays a major role in promoting healthcare IT, including EMR systems. The healthcare industry has adopted EMR systems widely. Nonetheless, nearly all EMR systems do not have an integrated electronic informed consent management system to support obtaining and automatically enforcing, patients' or participants' informed consents during treatment or medical research. In current practice, nearly all medical or research facilities use paper-based informed consent forms. The common approach used for managing informed consent forms in EMR systems is to scan paper consent forms and store them in electronic media, thereby restricting it to basic record keeping. Scanned consents are difficult to use by electronic information systems during medical processes, which makes caregivers responsible for enforcing the consents the same way it was done before introducing the EMR systems. Advanced automated features such as consistent enforcement of informed consent pursuant to statutes, regulations and guidelines are missing. In addition, hospitals spend on average $80,000 per year (Anonymous, 2007) scanning consent forms to perpetuate an inefficient, less than perfect consent system that results in subpar enforcement in the form of wrong or inapplicable consents. Failure to obtain informed consent is a top-ten reason for generating medical malpractice claims; "a properly completed consent form would have prevented 45% of errors on surgery wrong sites" (Compbell, 2004). Furthermore, studies concerning enforcement of medical informed consent in existing EMR systems reveal that no EMR systems support dynamically obtaining, and automatically enforcing, patients' informed consents when implementing medical processes. Consequently, a mechanism to implement and enforce e-consent is needed.

Informed consents in healthcare are governed by statutes and regulations that change over time. For example, a medical lawsuit that establishes precedent may trigger a new or amended statute or regulation or alter how a law is interpreted. Nevertheless, studies on consent laws are scarce, perhaps because ever changing laws would cause such studies to have a short shelf life. Issuing proper medical consent demands to evaluate all consent rules and to determine the type of consent to be obtained from the appropriate consent provider. Yet, no implementation method or modeling changes to consent laws are available that govern specific institutional practices (Beauchamp, & Childress, 1989). Consent is rendered unreliable without adhering to corresponding changes in laws. Moreover, consent statutes and regulations are complex. At times, ambiguous and complex laws (e.g., precedent changing interpretations of the law that, due to ambiguity, is subject to multiple interpretations) result in having gaps between user's understanding and the mechanized enforcements. Thus, a methodology for acquiring proper consent based on evolving law must be able to accommodate the dynamism inherent in changing regulations.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Authors' approach incorporates a medical consent management component into an existing EMR system to obtain and automatically enforce dynamically e-consent by using workflow technologies. More specifically, the authors use ontology-based knowledge representation and reasoning mechanisms to obtain required informed consents based on requester role, request purposes, each unique patient situation to ensure compliance with consent laws pertaining to the applicable sub-discipline. This technique also enforces protection of patient privacy in EMR systems. A case study, informed consent for genetic services, is used to show authors' work.

Following this Introduction, Section 2 describes informed consent in healthcare domain; Section 3 provides a case study, to show detailed models of authors' approach; Section 4 displays an implementation and, finally, Section 5, entitled "Conclusion," summarizes the proposed EMR system and briefly identifies future improvements.

MEDICAL INFORMED CONSENT

Categories of Consent

As mentioned above, medical informed consent has become the bedrock of medical practice. It majorly falls into three categories:

Treatment Consent

Patient safety, quality of care is the first major concern in practice medicine. Currently, every provider must obtain consent from a patient or guardian before commencing treatment, terminating treatment and choosing among alternative treatments. Nonetheless, the forms of treatment consent are various. In some cases, the patient or legal guardian must consent and/or sign a document memorializing that consent has been given. And yet, in most care providing entities, the patient showing up or requesting diagnosis or treatment constitutes default consent.

Informed treatment consent– either expressed or derived– is an important aspect of proper medical treatment, including but not limited to undertaking alternative treatments or terminating treatment. That process includes providing an unbiased risk/benefit analysis; explaining alternative treatments in a way that the patient understands, accurately communicating the care provider's understanding of the appropriateness of the treatment; advantages and disadvantages of obtaining the treatment or not obtaining it; and available alternatives (Appelbaum, & Thomas, 1995). The informed aspect of consent requires care receiver's awareness of he/she is willing to consent without undue influence by caregivers, insurers, or other involved actors.

Another requirement is caregiver acknowledgment that the patient and/or guardian have the mental capacity to consent. This makes psychiatric consents and consent given under anesthesia particularly difficult. Many federal, state and local governments have developed laws and regulations, and healthcare organizations have developed guidelines, for obtaining and memorializing informed treatment consent. Traditionally, the care seeker or his/her legal consent holder sign consent forms.

Much literature shows that researchers focus their research on different informed treatment consent requirements. For the standards of disclosing treatment benefits, risks, and alternatives, Studdert, D. M., et al. (2007) analyzed geographic variation in informed consent law. Authors in (Boonstra, & Nash,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2000; Bourgeois et al., 2008; Buchanan, 2004; Grisso, & Appelbaum, 1998; Holder, 1987) assessed patients' decisional capability to assent to treatment. In (Fisher, 2002), authors described four psycho-legal standards, communicating a choice, factual understanding, appreciation of the situation, and rational manipulation of information, all used to evaluate a patient's competence in giving consent. However, to date, this aspect of consent lacks a widely accepted standard.

Advances in medicine and technologies bring new challenges to treatment consent. For example, advantages in genetics being adopted as medically acceptable indicators to provide preventive and pre-emptive treatment for hereditary diseases. Gene therapy inserts genetic material into patients' targeted cells to treat disease. Up to date, gene therapy has been used to treat cardiovascular disease and cancer. Thomas Wirth (Wirth, Parker, & Ylä-Herttuala, 2013) made projections that "Gene therapy… will sooner or later be part of the standard care for a variety of different diseases". Unlike common treatment, informed consent delivers explicit risks/benefits of treatment to help patients in making a decision; it has been very difficult to determine precisely immediate benefits from gene therapy. Without full awareness patients' interests and choice may become compromised. Therefore, the contents of automatically generated consent forms should be dynamic, clear and explicit to the individuals.

Although less research has been devoted to managing treatment consent in EMR systems compared to protecting identifiable patient information, many researchers continue to improve consent treatment management. See for example, papers "Informed consent for clinical treatment" (Hall, Prochazka, & Fink, 2012) and "Informed Consent for Medical Treatment and Research: A Review" (Del Carmen, & Joffe, 2005) that introduced the basic concepts of informed treatment consent and made several practical suggestions for processing treatment consents effectively. However, none of those suggestions or analysis provides approaches to obtain dynamically or enforce automatically informed treatment consent in EMR systems. Even though, many healthcare organizations have attempted to include electronic consent management in their EMR systems. Veterans Administration Medical Centers use iMedConsent™ (VHA HANDBOOK 1004.05, 2005) that supports electronic access, completion, signing and storage of informed consent forms and advanced directives. iMedConsent has two parts: Software application and clinical content library. iMedConsent generates consents on each procedure without workflows, so the system is limited; it neither dynamically gains informed consents at the point of providing treatments nor enforces consents on medical procedures.

Some medical treatments are procedure-oriented, established by specialty boards as practice standards consisting of activities that caregivers must complete. Required informed treatment consents are embedded in some of those activities, which is an important factor that can change the medical treatment workflow. Standardized treatment procedures, documented in guidelines, usually require multiple checks. However, current EMR systems do not truly support and enforce pre-defined treatment workflow in run-time. To surmount this challenge, the authors proposed using workflow technologies to develop a Workflow-based EMR system in authors' previous work "Building Dialysis Workflows into EMRs" (Yu, & Wijesekera, 2013).

Authors' Workflow-based EMR system provides procedural choreography for treatment regimens using a Workflow Management System (WfMS). Authors' Workflow-based EMR system directly controls the progression of predefined and prebuilt care procedures that comply with local and national standards and policies; and also by providing flexibility that caregivers may need to continue treatment when, in their judgment, superior care compels moving to the next or alternate methods. Inability to obtain consent in emergencies is an example of an exception to a predefined workflow. Authors' system is also well suited to study healthcare delivery safety and quality care issues: Workflow logs perform an

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ongoing audit trail to satisfy accountability requirements and can also be used to improve the workflow. Furthermore, to enforce consent requirements, EMR systems must conditionally grant permissions upon prior receipt of valid, appropriate consent. Therefore, a Consent Management System (CMS) should be introduced in existing EMR systems. By leveraging the fact that medical consents accompany medical procedures, the best way to implement a CMS is by incorporating a medical consent management component into a Workflow-based EMR system. Without changing the Workflow-based EMR system architecture, the authors treat a CMS as an additional component of WfMS. Informed consents obtained by authors' CMS are stored in the EMR as part of patient medical records. The authors also designed additional user interfaces for users to issue missing but required treatment consents. See authors' paper "Consent-Based Workflow Control in EMRs" for more information (Yu, Wijesekera, & Costa, 2014).

Privacy Consent

Patient privacy is the second major concern in consents. In general medicine, informed privacy consent, provides the proper opportunity and knowledge for patients and/or research participants to understand and decide how the medical community can use and share their identifiable medical information. The Health Insurance Portability and Accountability Act (HIPAA) defines Protect Health Information (PHI) as "information that identifies the individual or for which there is a reasonable basis to believe can be used to identify the individual" (Federal Register, Vol. 78, No. 17, 2013, pp1689). HIPAA and other healthcare-related privacy laws mandate how patients' PHI may be used and disclosed. Every medical entity/provider must obtain a patient or guardian's informed consent before sharing and/or using patient PHI except for specified purposes of use and disclosure, such as treatment, for public health activities. In authors' previous work, the authors analyzed HIPAA in a great detail. However, HIPAA only governs covered entities.

Similar to treatment consent for genetic diseases, studies of, and methodologies for, protecting privacy of genetic information have lagged behind scientific advances. The harms of wrongly releasing or misusing genetic information are more than the physical harm, but so-called "psychosocial" harms (Strouse, 2005), which may cause societal discrimination (uninsurability, unemployability), even breakdown of family relationships. Prince et al. describe three practical genetic counseling cases that illustrate genetic discrimination (Prince, & Roche, 2014). The fundamental covenant of protecting patient privacy is embodied in patient-doctor privilege. Compromising genetic information privacy may stigmatize someone as carrying a genetic marker for a disorder or disease. Many researchers have suggested adopting traditional information protecting methodologies to protect patients' confidentiality. Yet, unlike other clinical datum, genetic information is unique and traceable to an individual or group of individuals who share a genetic inheritance (Mascalzoni, et al., 2008, & Gostin, Hodge, 1999). For example, certain genetic information of an individual may not only precisely identify him/her as high risk of certain hereditary disease(s), but also indicate that his/her relatives may have the same risks due to a hereditable gene. Consequently, traditional methodologies for privacy protection, such as de-identification, are not as well suited to the new area. Informed consent complied with proper privation laws tailored for genetic research and clinical usage constitutes a strong basis for ensuring appropriate genetic privacy.

In the past decade, consent management has received considerable attention from researchers, healthcare organizations, and publications proposing different ways to improve electronic consent management. Especially, most of those e-consent approaches focus mainly on sharing medical data, privacy, and security aspects (Grimson, Stephens, Jung, Grimson, Berry, & Pardon, 2001; Saranummi,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2005). For example, "eConsent: The Design and Implementation of Consumer Consent Mechanisms in an Electronic Environment," Coiera and Clark (2004) provided guidelines for designing an e-consent system. Another relevant work is by Ruan C. and Yeo S.S. (2009) that used a UML Model to design an e-consent system. They first identified various parts necessary to specify the e-Consent rules about patient record protection, then used UML to model the properties required by an e-consent system to make the associated patient record protection rules explicit and verifiable. That theoretical work was neither used to design nor implement an e-consent system that works with an EMR.

Rusello G. et al. (2008) proposed a consent-based healthcare management workflow where patients can control disclosing of their medical data based on workflows that are related to inter-institution transfers such as consults. Others have proposed e-consent management to be integrated with EMR or electronic health records systems (Bergmann, Bott, Pretschner, & Haux, 2007; O'Keefe, Greenfield, & Goodchild, 2005; Sheppard, Safavi-Naini, & Jafari, 2009; Win, & Fulcher, 2007). Win et al. (2002) in their paper "Implementing patients consent in electronic health record systems" expressed patient consent using an interface-based approach. Regarding enforcement of privacy consents in existing EMR systems, Bhatti, R. and Grandison, T. (2007) quoted several reports about healthcare privacy, concluding that IT system security mechanisms to enforce patients' consents for privacy were under-utilized and often bypassed. As mentioned in authors' previous works, privacy consent has a lot in common with treatment consent. Nevertheless, they are different in the following aspects:

1. **Issues/Objects are Dealt with Differently:** For privacy consent, patients decide how to use or disclose their identifiable medical information; by contrast, for treatment consent patients decide what should be done to their body.
2. **Who is Covered:** In privacy consent, HIPAA covers the medical information in what are known as covered entities. These are health providers/hospitals/health plans/healthcare clearing house and the like. Non-covered entities, such as private pharmaceutical factories have to abide by state privacy laws and organizational policies.
3. **The Type of Consent Affects the Government's Ability Override Patient Wishes:** The patient does not solely control consent regarding sharing of a patient's identifiable information. This type of consent also depends on the purposes of use and disclosure and a hierarchy of controls applies to the final decision. For example, HIPAA or state law can override a patient who rejects sharing his/her identifiable medical information that relates to a "Serious Threat to health or safety (OCR, 2003, pp 8)" Pursuant to an individual's right to be free of unwanted bodily invasion, patients have more power when deciding their treatment procedures. Consequently in normal care, providers must execute valid treatment consents.

The differences between the two categories of informed consents preclude using the same methodology to obtain informed consent for protecting privacy, especially because privacy rules that reflect public policy may override patient wishes. Thus, the additional factors of governmental override add complexity to protecting privacy for information sharing management in EMR systems. In HIPAA, the purpose of using and disclosing PHI provides a good example; it requires additional components for managing and enforcing this kind of consent in EMR systems in different levels. The traditional authentication strategy, like Role-based Access Control (RBAC), is insufficient (Al Faresi, Wijesekera, & Moidu, 2010). Authors' approach combines patient consents with the hierarchy privacy rules of regulation to determine the correct subsequent actions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Participating Research Consent

Protecting and respecting participants of human subject research is the third major concern in healthcare domain. In most cases, informed consent for participating in human subjects research is a prerequisite. Participation in research allows researchers to generate and analyze new data or use their old data/samples. Traditional medical research may be well designed within a fixed time. The participants are informed with sufficient information about the study purpose, the procedures used in the study, and potential risks and benefits, etc. and the participants have the right to withdraw their consent at any time.

Contemporary medical research, especially biomedical research, provides many challenges to the currently accepted ethical standards of informed consent in both categories (Mascalzoni, et al. 2005; Knoppers, 2005; Petersen, 2007)

One challenge of designing consent for participating biomedical research and consent for performing a medical treatment is: in general medicine, clinicians seek patients' permission to perform planned procedures and/or to use their medical information for treatment. The patients are knowledgeable about the risks/benefits of performing procedures. Correspondingly, the benefits or risks of participating a genetic study may not be foreseen at the time of consent. As stated in (Strouse, 2005; The Presidential Commission for the Study of Bioethical Issues, 2013), obtaining informed consent in biomedical research settings present differences in required techniques and approaches when compared to clinical practice. In addition, informed consent is a process rather than an occurrence at the one-point action. Another challenge of consent for reusing sensitive medical data or samples in research is: in general medicine, patients reasonably consider that their medical information constitutes their own private information and, hence, are unwilling to either share or have divulged such information beyond participated study, as well as they, are easily able to withdraw such consent at any time. Conversely, the new genetic researches based on the enormous existing data or samples around the world, and may develop over time, which cause re-consent or completely strike out widely shared genetic data is impossible. The fundamental ethical principle of research consent may no longer be suitable. Irrespective of the purpose of research, methodologies for respecting participants and preventing them from both emotional distress and social risk have not, however, kept pace with rapid medicine and technique advances.

Hence, it is no surprise that some genetic medical practices and biomedical research are performed without obtaining appropriate informed consent. For example, researchers may entice participants to participate in a proposed study without providing them with the proper informed consent forms. To address this issue, some researchers advocate different methodologies such as using overly stringent policies to protect participants from harm, and to encourage trust in the research study itself. Obtaining written informed consent of testing to produce genetic data is such an example. But this approach potentially risks limiting scientific innovation (Hallinan, et al., 2015) Genetic research advances also benefit public health (Khoury, & Dorman, 1998). One of the evidences is breast cancer 1 (BRCA 1), where a gene mutation that can be used for general population to predict and prevent from disease (Struewing, & Abeliovich, 1995). Other researchers have proposed a new, open-consent model for medical and scientific genetic research (Lunshof, et al., 2008) or open-access policies for genetic data sharing (Kaye, et al., 2009) in order to achieve research goals quickly. Some researchers disagree on these proposed divergent methodologies and question if such policies may result in a new threat to research participants. Without the consensus of ethical and legal framework for biomedical research consent may result in a perceived breach of the patient's right to privacy and erosion of the ever-sacred principle of researcher/doctor-patient confidentiality (Petronio, 2012; McGuire & Secko, 2009; Beskow, 2010). The authors

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

balance the power of research goads and individual interests by implementing and enforcing related laws and organizations guideline in authors' designed prototype consent management system, for example, by recommending the conditions to apply a waived research consent by the federal "common rule" in federally funded research; guidelines for broad consent given by the Council for International Organizations of Medical Sciences (CIOMS, 2006). Additionally, the nature of genetic research may not promise short-term benefits to the participants and considerable risks in the research must be as part of participating research consent forms.

Components of Consent Forms

Due to the absence of an agreed upon views on the components of medical consent forms, the contents of consent forms may vary greatly. Nonetheless, many studies explored this issue and proposed various recommendations on the essential content of informed consent. Carmen Ayuso et al. (2013) proposed ten minimum elements of information to be included in all informed consent forms for whole-genome sequencing (WGS) in the clinical setting. They did so after searching and reviewing applied documents from 554 documents in the Pubmed-Medline and Embase, such as Scope, Benefits, Risks, Voluntary, Alternative test, Future use, Incidental findings, etc. The authors also examined many paper-based consent forms and found that they include a large number of attributes. As shown in authors' previous work, the authors found 21 attributes in eight dialysis consent forms. Authors' examinations have led to categorizing the consents as follows:

1. **Informed Consent Giver:** The person with the legal right to make health care or participating decisions, such as parents or legal guardians of minors, healthcare proxies, healthcare providers or third parties.
2. **Disclosed Information:** At a minimum, includes procedure (treatment procedure for treatment consent, data usage for privacy consent, research procedure for research consent, and so on), benefits, risks and alternative procedures.
3. **Patient's Decision:** Includes the decision (deny or accept) by providing all required conditions and attributes such as signatures, date, etc.

Consequently, consents are based on all input variables, such as consent giver, treatment specification, related consent rules, and the like. Usually, they are represented in the physical forms in writing, electronic forms in the computer applications, or in some other legally acceptable manner. Hence a well-designed consent form should be obtained when activities require it. In authors' prototype system, electronic consent forms are automatically generated based on the collected attributes and regulations and enforced to obtained as practice workflow demands.

The various rules generated from laws, organization policies or sub-discipline medical guidelines affect each component of consent. Therefore, to be valid, informed consent must comply with state and federal laws and regulations applicable to medical sub-disciplines, as well as depend on some specific situations such as treatment type, research type, facility location and the purposes of use and disclosure of sensitive patient information for medical information sharing consent.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Consent Rules

As stated above, the relationship between consent components and governed rules are expounded below:

1. Informed consent giver component is governed by a competency rule; that is there to validate that the patient has sufficient maturity to grant informed consent;
2. Disclosed information component is governed by information or disclosure rules. An essential component is the concept of autonomy that allows competent adult persons and emancipated children to make their own decisions based on the understanding of given information. Nonetheless, the scope of the disclosure of required information is still debated, especially in genetic consents.
3. Patient's decisional component is governed by decisional capability rule, which is for evaluation of patient competency for understanding information about the procedure or using/sharing medical data and the ability to provide a rational, voluntary agree or deny decision. In (Fisher, 2002), authors described four psycho-legal standards of:
    a. Communicating a choice,
    b. Factual understanding,
    c. Appreciation of the situation, and
    d. Rational manipulation of information that is used to evaluate patient competence in granting consent.

However, no widespread consensus exists on how the decisional capacity of treatment consent should be measured.

Consent statutes and regulations are complex. Authors' previous work (Yu, Wijesekera, & Costa, 2014) showed the complexity of determining patient competency for treatment consent. One of the issues is that each state has its own laws defining minors, at times differing by treatment type. Various statutes and regulations and laws governing informed consents change over time. For example, a medical lawsuit that establishes precedent may trigger a new or amended statute or regulation or alter how a law is interpreted. Nevertheless, studies of consent laws are scarce, perhaps because ever changing laws would cause such studies to have a short shelf-life. Yet, no implementation method or modeling changes to consent laws are available that govern specific institutional practices (Beauchamp, & Childress, 1989). Consent is rendered unreliable without adhering to corresponding changes in laws. Even more, at times ambiguity adds to complexity (e.g., precedent changing interpretations of the law that, due to ambiguity, is subject to multiple interpretations). This causes breakdowns between concepts that people use and data that computers interpret. Thus, a methodology for acquiring proper consent based on evolving law must be dynamic.

Issuing proper medical consent demands to evaluate all consent rules and to determine the type of consent to be obtained from the appropriate consent provider as well as allowing exceptions to achieve the flexibility. For example, an incapacitated patient (perhaps temporarily due to injury or medication) may be considered not to possess a sound mind or may have been determined not to have the mental capacity (such as developmental disorders or clinically determined mental impediments) needed to provide legally binding informed consent. Similar exceptions exist with respect to children or young adults, who may not be considered mature to provide consent (Appelbaum, & Grisso, 1995).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

System Architecture

Authors' system architecture (shown in Figure 3) consists of the following components: (1) User Interface (UI) for EMR system users, (2) EMR's Runtime System, (3) WfMS ─ Workflow Management System - a runtime system that enforces medical treatment workflow, (4) CMS ─ Consent Management System, and (5) corresponding databases.

1. **UI:** Enables interactions between EMR systems and EMR system users (usually refer to people, such as members of the treatment team and quality care team that may want to review the logs, etc.)
2. **EMR System:** Provides clinical functions, including the display of patient demographic information, vital signs, medication, lab order/results, etc.
3. **Workflow Management System:** Designs and executes workflow models, consisting of a workflow editor, workflow runtime engine and other components such as a workflow task handler that enforce the completion of a specific task.
4. **Consent Management System:** Evaluates and obtains appropriate patient's consent form, if needed. It contains different consent services, such as Patient Privacy Consent Management Service (PPCMS) and Patient Treatment Consent Management Service (PTCMS).
5. **Related Databases:** WfMS' databases contain healthcare and medical research workflow specifications composed of tasks describing processes and WfMS' data; while EMR systems' databases contain patients' or research participants' medical records and other EMR systems' data.

As stated, previous work describes the details of the developed consent management and workflow-based EMR design. The enhanced Workflow-based EMR system includes:

A CMS connected to WfMS that contains different consent management services, such as treatment consent management service, privacy consent management service, participating research management service, and so forth. The authors describe two of main consent management services below:

1. A PTCMS that contains a Treatment Consent Management System (TCMS) and Treatment Consent Rule Management System (TCRMS), which ensure treatment consents comply with treatment consent laws and policies. TCMS generates the appropriate informed treatment consent forms automatically and obtains the consent before passing to the next task.
2. A PPCMS contains a Privacy Override Privilege Management System (PoPMS) and a Privacy Consent Rule Management System (PCRMS). When an entity requests using or disclosing protectable patient information through the EMR system, the sub-component, the PoPMS in the system will check the requester, which may be a government agency, patient, covered entity, business associate or a researcher. Then PoPMS retrieves the purpose of disclosure and use of sensitive medical information based on federal, state privacy rules or organization policies. For example, in HIPAA, without any stated purposes, a patient or his/her authorized representatives is allowed to require this patient's PHI. In this situation, the request is assigned as "required disclosure" by default. If the request is an "authorized use" and disclosures, the request will trigger PCRMS to obtain a patient's informed privacy consent, which when given by the patient can be used for disclosure. In this design, the Ontology Service (OS) plays a role in obtaining valid consent forms. OS includes an ontology editor (to edit terms and rules), a rule set that codifies treatment or privacy consent

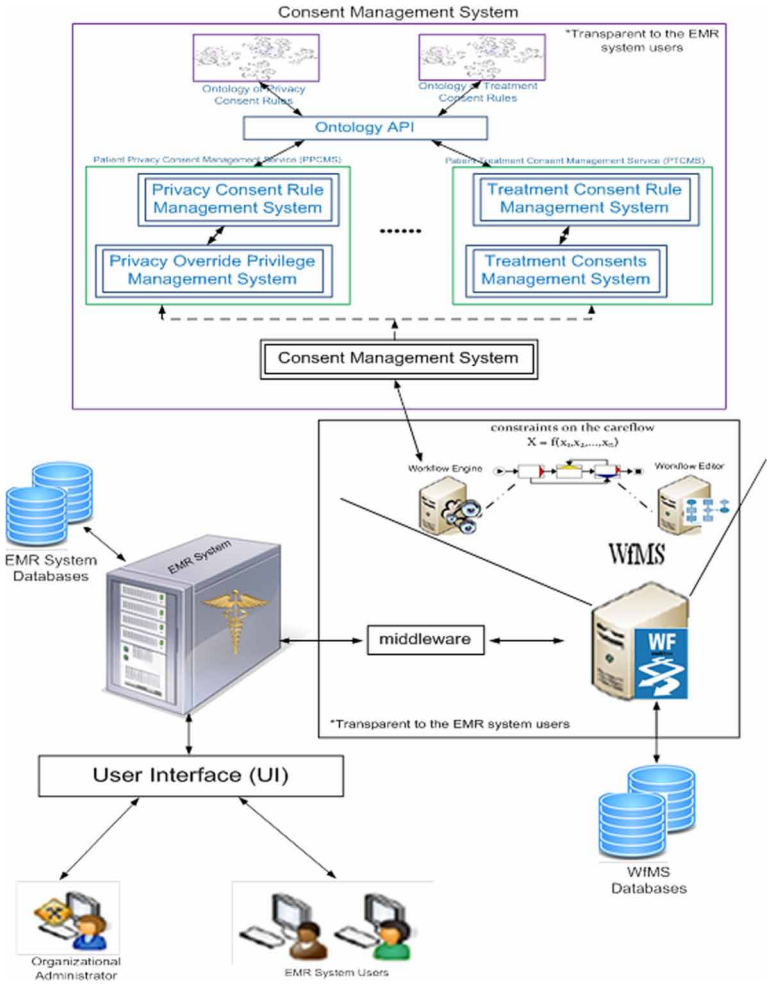
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

laws and a reasoner that derives components of the consents that are displayed as consent forms by the EMR system. PCMS returns the final decision made by privacy rules with the highest override privilege to the EMR system. For authorized uses and disclosures request, the decision is shown on a patient's valid privacy consent.

Similar in design to the TCMS, PoPMS can retrieve or generate the valid informed privacy consent forms automatically and passes the patient's decision of sharing his/her sensitive medical information. The Enhanced Workflow-based EMR System Architecture is shown in Figure 1.

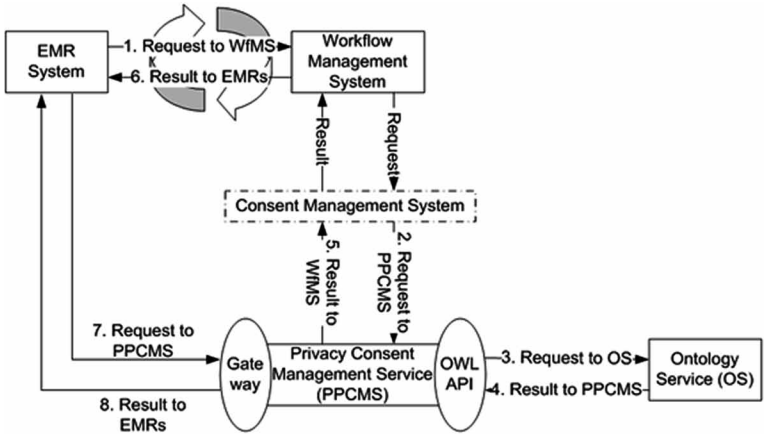
Informed consent is managed by the combination of WfMS and CMS. Based on the kind of consent required, corresponding consent management service is triggered. Following, the authors use PPCMS as an example to show how does the system work. PCRMS must be able to ascertain the consents, if any, are missing, but are required for making the final decision; if so, they must be obtained prior to disclosure. See, Figure 2 for the interaction among the WfMS engine; the EMR systems; and PPCMS and OS.

Figure 1. Workflow-based EMR system with consent management architecture



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Interactions between the WfMS; the EMR system; and PPCMS and OS per session



Step 1 EMR System → WfMS: When a requestor starts to request an individual's sensitive medical information in an EMR system or the request is issued during other workflow tasks, a "launch case" event request with workflow specification I.D. or name is sent to WfMS engine; WfMS engine enables the work item(s) based on control flow defined in the workflow specification and input data. If the enabled work item(s) do not request consent service, then go to Step 8. Otherwise,

Step 2 WfMS → CMS: If a procedure/step needs to check patient's informed privacy consent, the CMS is triggered. Furthermore, base on requested consent type, a corresponding consent management service is triggered, PPCMS in this case.

Step 3 PPCMS → OS: PPCMS uses OWL API to connect to the OS with patient's information and other required consent information.

Step 4 OS → PPCMS: OS retunes the results reasoned based on the rules to the PPCMS.

Step 5 EMR System → PPCMS: This additional step is triggered only when informed consent is required. EMR system requests the PPCMS to the kind of the privacy consents should be issued.

Step 6 PPCMS → EMR: System Similar to the previous step, except this step is triggered only if the consent service is called. PPCMS returns the answers to EMR System.

Step 7 PPCMS → WfMS: PPCMS passes results to CMS, then the CMS forwards results to WfMS, the WfMS decides if procedures should continue or be aborted based on the workflow specification.

Step 8 WfMS → EMR System: WfMS sends a notification to EMR system. The available work item(s)/ step(s) will notify EMR system users via UI. Then, the interactions between WfMS and EMR system are repeated.

In the next section, the authors use informed consent for genetic services as a case study to show how genetic informed consents placed new challenges on the traditional ethical standards of informed consent. It also illustrates how appropriate consents can be electronically obtained and automatically enforced using a system that combines medical workflows and hierarchically, ontologically motivated rule enforcement.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CASE STUDY ON GENETIC SERVICES

In authors' previous works (Yu, Wijesekera, & Costa, 2014), the authors mainly address the consent for medical treatments, with a focus on informed consent for procedure-oriented treatment regimes. A preliminary study convened jointly by the National Institute of Health (NIH) and the Center for Disease Control (CDC) (Clayton, 1995) presents the risks and ethical issues involved in collecting and storing tissue samples for genetic tests in a research setting. Herein, the authors present a unique approach, which is based on the premise that acceptable genetic service regimes are captured in workflows to automatically obtain or generate (when needed) the deferent types of informed consent for clinical usage (collectively called *genetic services*) and genetic research. Moreover, the authors combine genetic privacy laws and policies relevant to informed consent with an encompassing and customizable genetic services workflow to show consent enforcement for ensuring appropriate genetic information privacy.

Background

Genetic studies comprehensively utilize genotypic and phenotypic data (Ritchie, et al., 2015). Studies have shown that preventive care, in which genetic services and researches play an important role, costs significantly less than treatment upon disease onset (Németh, et al., 2013). Examples include genetic mutations causing certain kinds of breast cancers (BRC-1 and BRC-2), ovarian cancer, sickle cell anemia, behavior disorder, β-thalassemia, left ventricular noncompaction cardiomyopathy and Alzheimer's disease, etc. Those genetic advances have encouraged some genetic services to become medically acceptable common practice to provide preventive and preemptive treatment for hereditary diseases. Concurrently, rapid advancement of genetic research continues to lengthen the list of predictable diseases (Rehm, et al., 2016). Many of those achievements are produced by researches using large sets of stored genetic information and samples. As stated in (Kerath, et al., 2013), the utilization of existing biobanks' stored information has become the cornerstone of human genomic and epigenomic studies. The concept of open access in the field of genomic research is expressed in the Bermuda Principles and the Fort Lauderdale Agreement, which have been widely applied in North America and the UK for funded genetic research. Sharing genetic data is vital for maximizing the use of such information to benefit patients and the general public. However, concern for ensuring genetic privacy is growing.

Genetic exceptionalism is defined in a paper (Gostin, & Hodge 1999) as "*the societal practice of treating genetic data as different from other types of health data for the purposes of assessing privacy and security protections*." The Federal Government and some states enacted several genetic-specific legislations to protect the genetic privacy and prevent discrimination. For example, HIPAA considers genetic information to be confidential medical information and regulates healthcare providers. In spite of Gostin & Hodge (1999) argued this definition is flawed in the light of two reasons:

1. Strict protection of genetic data threatens the accomplishment of public goods;
2. There is no clear demarcation separating genetic data from other medical data.

The majority of scholars agree that genetic information has its particular characteristics from other medical data. Some researchers say that *genetic information is essentially familial in nature and is referred to as the Genetic Information is Familial Thesis (GIFT)* (Liao, 2009), sharing such information will benefit related groups of individuals. Yet, others, who advocate the view of humans possessing the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

rights of *do not want to know* (Sherlock, 2009, & Liao, 2009), disagreed with the familial approach in their literatures. Sharing such data introduced new challenges. Inter alia, related regulations are many. For example, some countries have regulations to enforce sharing such information among family members (Lucassen, & Kaye, 2006, & American Society of Human Genetics Social Issues Subcommittee on Familial Disclosure, 1998) and others do not. Both research and clinical uses of genetic information pose ethical issues on medical consent:

- **Ethics:** Privacy involving genetic information differs from traditional medical informed consent. For example, protecting patients' private information is an important medical ethics and legal obligation. The key point is that the healthcare industry's current level of privacy and confidentiality is not yet entirely suitable for genetic privacy in both clinical medical practice and biomedical research. Genetic privacy consent has not been linked-up to applicable laws as part of an EMR workflow that enforces appropriate consent as a predicate to using genetic information in research and clinical practice. Protecting patient privacy in the context of genetic information leads to other equally unique types of challenges.
- **Legal Issues:** Requirements differ based on the usage. For example, generating data requires written informed consent that meets strict requirements, whereas sharing genetic data for research may have a less stringent opt-out consent. Genetic information can be traceable not only to individuals but also identify data owner's relatives and makes family trees susceptible to stigma (Roach et al., 2010). Therefore, genetic privacy has been proposed as categorical privacy that differs from traditional individual-centered concepts of privacy in literature (Lunshof, 2008). Furthermore, genetic privacy laws vary from state-to-state and may be inconsistent with and be more or less stringent than federal law.
- **Social Implications:** Societal views are often reflected in law and/or organizations' policies, so societal implications are likely inextricably intertwined with laws and policy governing genetic privacy and what constitutes informed consent.

As a solution, the authors provide an encompassing framework consisting of workflow-enforced genetic privacy as well as biomedical consent management consistent with state and federal genetic privacy laws such as a statute, regulation, and precedent. These missing regulations presented issues for the authors' prototype implementation that have been resolved by having default rules.

Related Work in Context

EMRs play a vital role in sharing medical information among participating actors based on their usage scenarios. In addition, using EMRs for genetic services present a unique set of challenges (Mascalzoni, 2008). Belmont et al. (2009) highlighted the privacy, ethical and legal issues of handling genetic data in EMRs. Additional, Scheuner et al. (2009) conducted a case study to verify if existing EMR systems meet genetic information needs. This study, involved results, and conclusions gathered from discussing 56 patient's electronic medical records within10 EMR specialists, 16 medical geneticists, and 12 genetic counselors, shows an overall lack of support for functionality, structure, and tools for clinical genetic practice. A more recent study on the state of EMRs supporting genomics for personalized medicine again identifies the designing of the structure of data as a challenge (Ullman-Cullere, 2011). The authors also identified genetic services workflow management as a priority area that needs further

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

research, development, and testing. The authors, too, observed that current EMR systems for genetic services still cannot support the ability to obtain/generate informed privacy consents and enforce them in EMRs (Yu et al., 2015).

Some researchers suggested that the proper legislation on generating and using genetic information is pivotal to improving genetic privacy (Gymrek et al., 2013). In 2013, HIPAA (45 C.F.R. §160.103) Omnibus Rule included genetic information as PHI to be regulated under the privacy portion of HIPAA. Nonetheless, states may have different definitions of genetic information. The combination of Federal privacy laws and state laws from many states form a fragmented regulatory and statutory landscape for permissible information sharing and consent management. Indeed, significant regulatory gaps create additional burdens in providing automated ways to obtain and generate information consent in EMRs.

As mentioned, an informed consent management system in genetics is needed. To the best of authors' knowledge, to date, there is no such system implemented in EMR systems that can accommodate the rapid advancements of genetics research and the evolving nature of the genetic privacy landscape. To comply with different genetic privacy laws under various circumstances in clinical practice and genetic research is challenging. EMR systems evolved with differing architectures to satisfy different requirements at different stages of technology development. Many treatises (Locatelli, Restifo, Gastaldi & Corso, 2012; Shortliffe, & Cimino, 2006; Wickramasinghe, & Geisler, 2008) have detailed EMR system evolution. Today, EMR systems are indispensable. Because medical organizations have over two decades adopted costly EMR systems ($25M-$45M per large hospital and $8B for the Defense Health Agency), medical organizations are understandably reluctant to switch to new systems. Consequently, a strategy of integrating a genetic consent management with existing EMR systems may provide a solution that the authors advocate.

Proposed Approach

Authors' solution to the issue is to enhance an existing EMR system with a workflow management system to choreograph predefined procedures, as the authors have prototyped using open source solutions. Although there will be other components that drive the EMR systems, EMR users will not notice a difference in interaction by using authors' solution.

During the time that genetic services have been around sub-disciplines or education papers have also defined de-facto workflows for some genetic services. For example, The Genetic Counseling Definition Task Force defines Genetic counseling as "The process of helping people understand and adapt to medical, psychological, and familial implications of genetic contributions to disease." However, the authors were unable to find a documented and modeled comprehensive workflow. Therefore, the authors need to create a genetic services workflow model. The authors describe a genetic services workflow specification in the next section.

High-Level Genetic Services Workflow

The methodology used to generate a genetic services workflow involves the following steps:

1. The authors interviewed genetic researchers to get their individual practice workflows.
2. The authors collected the paper-based documents used to record events and data that are associated with the workflows.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

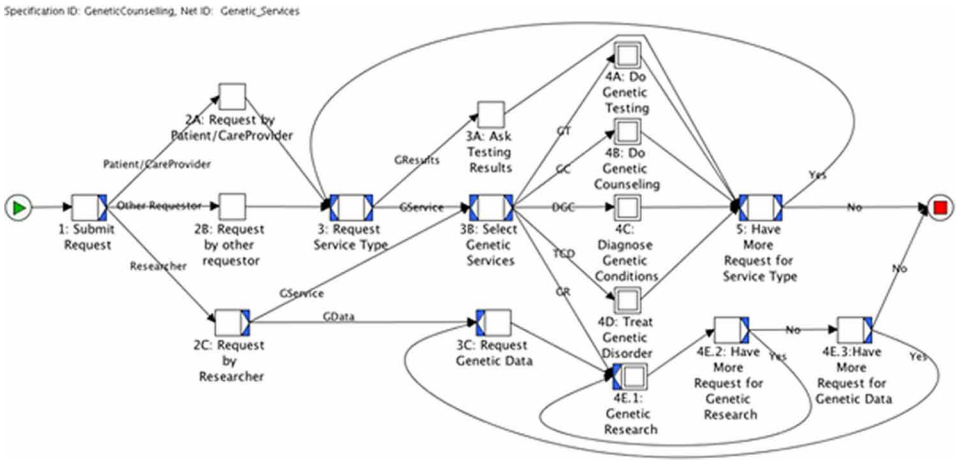
Finally, the authors combined (1) and (2) to generate a more comprehensive workflow. The authors describe the high-level workflow process first; then describe its one of sub-processes genetic counseling with the requesters involved in the processes separately and show an embedded consent management system.

The high-level genetic services, modeled using a workflow editor is shown in Figure 3. The shown workflow covers the high-level view of the genetic services consisting of multiple requested genetic services, as shown in Figure 1, the leftmost node with a right-pointing green arrow inside a circle is the starting point of the workflow, and the rightmost circle with a filled red square ends the workflow. Atomic Task (using a single box in Figure 3, as (1: Submit Request) denotes "a single task to be performed by a human participant or an external application"; Composite Task (using a double box in Figure 1, as (4B: Do Genetic Counseling) represents "A container for another workflow process (also called a sub Net)"; The starting point is the process of receiving different request for a genetic service by different requestors, such as care providers, researcher, etc. When genetic service or genetic information using/sharing, the consent management system is triggered automatically in authors' design. Details follow.

"Do Genetic Counseling" Sub-Workflow

The continuous workflow that the authors call "do genetic counseling," named *4B: Do Genetic Counseling,* has two composite tasks in *Do Genetic Counseling* sub-workflow, *Pre-test Genetic Counseling,* and *Post-test Genetic Counseling,* as shown in Figure 4 and Figure 5. This sub-workflow is based on a presentation on clinical genetic counseling from Stanford University's Clinical Genomics program (Merker, 2014). In addition to the previous presentation, the authors used another presentation by O'Daniel et al. (2012) on Genomic Medicine for additional details. Here the authors provide a detailed documentation and overview of the Genetic Counseling workflow. The following list provides the details of the individual tasks involved in the process of providing genetic counseling to a patient or care provider.

Figure 3. High-level genetic services workflow



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Pre-test genetic counseling

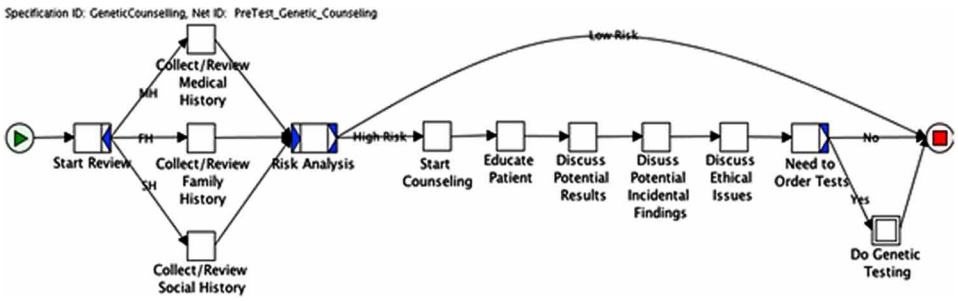
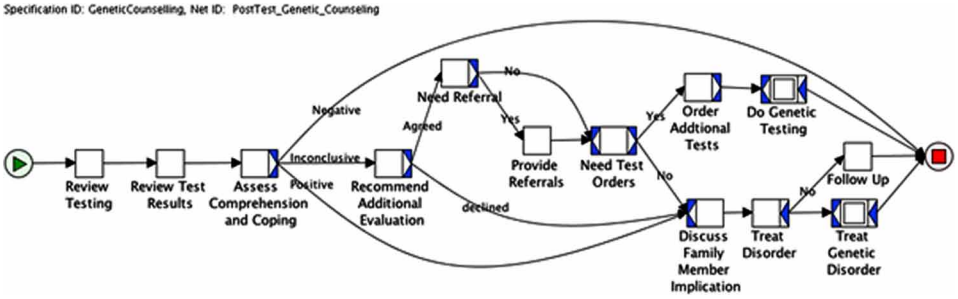


Figure 5. Post-test genetic counseling



1. Collection of Medical Records, Family History, and Social History:
   a. **Require Genetic Counseling:** A case originates with a physician's referral for genetic counseling. Alternatively, a patient self-requests genetic counseling after learning of a manifestation of a genetic disorder in a family member.
   b. **Patient Walk-ins:** Upon arriving at the genetic counselor's office, a self-requests patient is asked to provide social and family history. Concurrently, the Genetic Privacy Consent is obtained. Additionally, if copies of medical records are provided, they are recorded in the system. In authors' workflow model, the patient is presented with a detailed family and social history questionnaire to assess the risk of a suspected hereditary disease. Authors' workflow produces a questionnaire on suspected Hereditary Cancer Syndrome based on forms developed by the Virginia Women's Center (Virginia women's center website).
   c. **Physician's Referrals:** When a physician refers a patient for genetic counseling, the patient is asked to sign release forms for medical records (Genetic Privacy Consent) from the referring physician (if not already provided). If not, and the patient brings in copies of medical records, they are recorded in the system.
   d. **Review Requirements:** Information collected from the patient is reviewed for completeness. If any, missing information is collected from the patient, patient's primary care physician or public sources of ancestry information. This is shown in the "*Collect More Information*" task in the modeled workflow (Figure 4). When all of the information (to the best of patient's knowledge) is collected, one or more of following people review the collected information: Genetic Counselor, Molecular Pathologist, and/or Medical Geneticist. Next, check for the amiability of genetic test to answer the questions posed by the patient or the referring physician.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. **Pre-Test Counseling:** Once an insurance or payment authorization is obtained, the patient meets with a genetic counselor for pre-test counseling. The genetic counselor performs the following sub-tasks for pre-test counseling:
    a. An informed consent is obtained for counseling, i.e., Genetic Counseling Treatment Consent. As a part of this informed treatment consent, the counselor provides the patient with information about the expected risks and benefits of this process.
    b. The family, social, and medical history provided by the patient is reviewed.
    c. The counselor discusses the expected range of results and their impact on the patient and his/her relatives. Additionally, the counselor discusses other potential incidental findings that may be in the report. Incidental findings are those findings in the test results that are not associated with the primary reason for conducting the test, but indicative of other recognizable genetic disorders. The counselor also discusses ethical issues that could arise from conducting or disclosing these test results. Before ordering the tests, any questions or clarifications from the patient are answered. At this moment, the counselor will obtain a doing lab consent, which belongs to the Medical Treatment Consent.
    d. If applicable, consent for sharing information or using their samples for current or future possible research purposes (Participating Genetic Research Consent) is collected.
3. **Lab Work:** Blood or tissue samples are collected from the patient for genome sequencing. Generally, one tube of blood is collected for sequencing. An additional tube may be collected for confirmatory studies. Another tube of blood may be collected from the patient with appropriate consent for research purposes. The blood and tissue samples are then sent to a laboratory for sequencing, and the sequenced genetic information is returned to the genetic counseling team. The authors do not model the workflow for genome sequencing in the laboratory in this chapter.
4. **Post-Test Counseling:** Once the sequencing results are reported back to the genetic counseling team, a draft report is prepared. This draft report is reviewed by a review team consisting of Genetic Counselors and/or Medical Geneticists, Physicians, and/or content experts. Once reviewed, the final report is uploaded into the EMR and is ready for the patient. Ideally, this report should be delivered electronically to the EMR. When the patient meets the counselor to discuss the test results, the following sub-tasks are performed:
    a. The counselor reviews the test results with the patient. Once the review is complete, the counselor assesses the comprehension and coping skills of the patient.
    b. If the additional evaluation or referral for treatment (e.g. prophylactic mastectomy) is warranted, referrals or orders are placed by the genetic counselors. Alternatively, the patient may decline the recommendations of the genetic counselor. If the recommendations are declined, it is recorded in the EMR and if needed, a second opinion is scheduled.
    c. The counselor finally discusses implications of the findings on related family members. At this point, the genetic counselor may recommend the patient to inform related family members personally about the test results and treatment options by providing templates. Alternatively, they may obtain consent (Genetic Privacy Consent) to disclose test results to affected family members.
    d. Lastly, if required, the counselor may schedule follow-up appointments.
5. **Disclosure and Other Consents:** At the end of the genetic counseling process, the care provider may be required to obtain consent to retain samples and/or genetic information (sequencing results) and genetic test reports based on state and federal regulation. Additionally, the care provider may

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

require additional disclosure consents to share information with other individuals or entities as required by law. If consent to retain genetic samples and test information is denied, they must be destroyed and/or purged from the EMR records. Once these housekeeping tasks are complete, the genetic counseling workflow comes to an end.

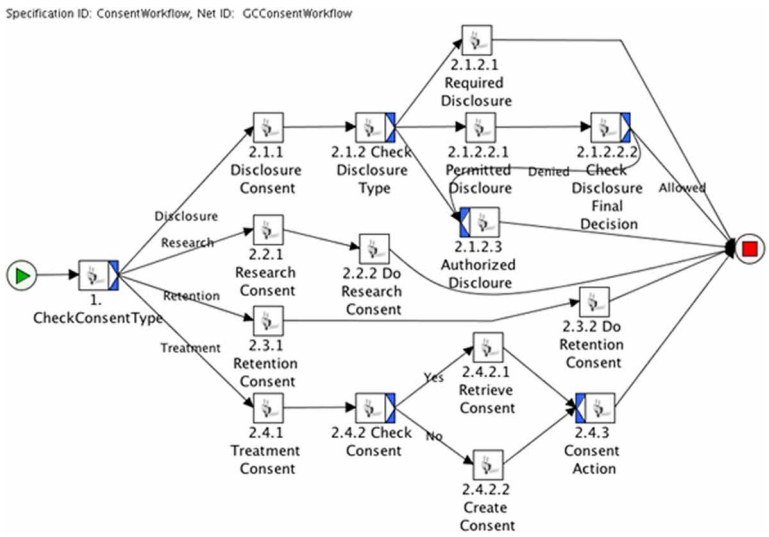
Consent Management in the Genetic Counseling Workflow

As discussed earlier, there are various types of informed consent required for Genetic Counseling, which mainly include treatment consent; genetic information use/disclosure consent, so-called genetic privacy consent; genetic research consent; and retention consent. If a proper workflow in not enforced for genetic counseling in the EMRs, it may leave healthcare providers and/or health-care organizations open to lawsuits. Existing laws limiting genetic information disclosures do not generally consider ethical issues involved during the disclosure process. Ethical issues present an abstract concept that needs to be carefully considered by genetic counselors on a case-by-case basis. Additionally, an organization's ethical guidelines may regulate disclosure of genetic information, which upon disclosure may tangibly affect a third party.

In this section, the authors describe how the authors enforce all of the required consents in the genetic counseling workflow. The authors describe the enforcement of disclosure consent in this work. Other types of consent may be enforced in the workflow very similar to the disclosure consent.

When a task requires consent in the genetic counseling workflow, the consent management system is invoked to manage the consent requirements. The authors also adopted workflow technologies to model the consent management processes as a workflow specification, shown in Figure 6. In the case of information disclosure consent, the consent must comply with federal and state laws, as well as organization policies.

Figure 6. Genetic consent management workflow



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In developing authors' prototype, the authors studied genetic-related privacy laws in great detail. At first, the authors examined the papers and publications (Rothstein, 2014; & American Health Information Management Association, 2013), which focused on the federal genetic privacy laws. There are four major federal laws for genetic information sharing as shown below:

1. Genetic Information Nondiscrimination Act of 2008 (GINA) (2008). GINA regulates employers and health insurance companies but not healthcare providers who use genetic information. GINA protects individuals from discrimination based on genetic conditions. However, GINA does not apply to federal government employees or employers with fewer than 15 employees.
2. HIPPA (45 C.F.R. §160.103). The 2013 Omnibus Rule included genetic information as PHI to be regulated under the privacy portion of HIPAA. However, HIPAA only mandates a portion of genetic information requestors such as covered entities.
3. The Common Rule applies only to institutions that perform federally funded research on human subjects; private research institutions may voluntarily agree to comply with federal standards. Among other issues, the Common Rule sets out explicit standards for informed consent by research subjects, although an ethics board may waive these requirements.

Although these laws guide establishing policies for genetic information sharing, none of them absolutely protects genetic privacy. In addition to federal laws, many state laws and institutional polices add to the informed consent management.

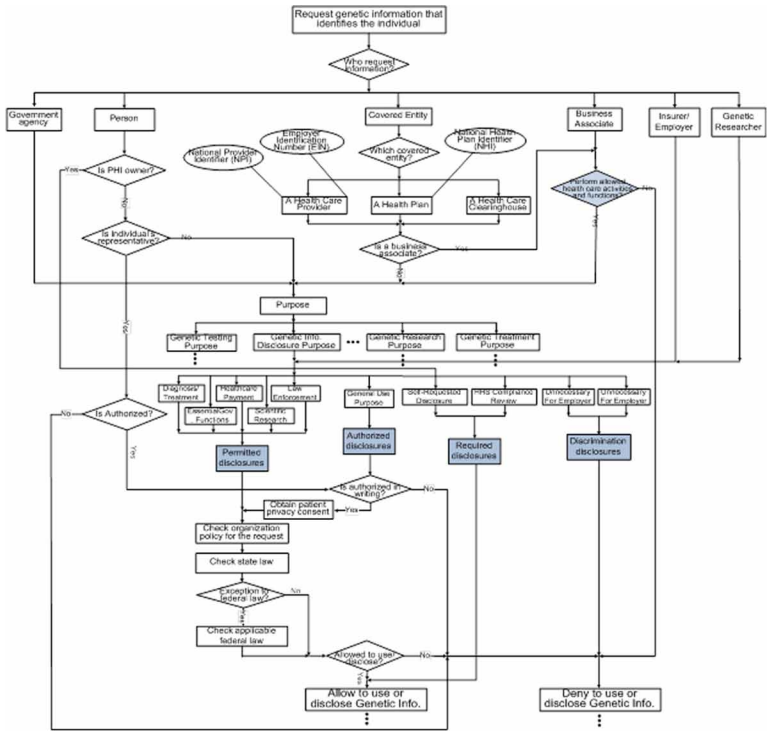
Now the authors are reviewing genetic privacy laws from states that apply to genetic privacy in published literature (Mulholland, 1998; & National Conference of State Legislatures, 2008) and online resources. Many states have their own genetic privacy laws with different levels of requirements to protect genetic privacy.

For example, California has a more strict genetic privacy law while some states such as Alabama do not have any regulations specific to genetic data. Furthermore, the genetic information use and sharing of genetic services will vary based on the types of the requestor (such as genetic counselors, researchers), and the purpose (such as diagnosis, research, payment). Without an understanding of informed consent, it is problematic to make consent information semantics explicit and hence discoverable automatically in a software system.

Computer science defines "Ontology" as "the explicit specification of a conceptualization" (Gruber, 1995) or "explicit formal specifications of the terms in the domain and relations among them (Noy, & McGuinness, 2001)." Consequently, ontologies can be used to model a domain of discourse as a set of representative primitives, such as classes, object properties, data properties and relationships between them. Many papers present reasons to use ontologies to formulate shared vocabulary (knowledge) of a given healthcare domain of interest (Gruber, 1993), to ensure interoperability within the various healthcare domains (by ontology matching). Ontologies have been used to represent actionable knowledge in biomedicine (Dang, Hedayati, Hampel, & Toklu, 2008; Garde, Knaup, Hovenga, & Heard, 2007; Matos, Campos, Braga, Palazzi, 2010; Rosse, & Mejino Jr, 2003; Yu, 2006) due to their ability to capture domain knowledge in a formal yet simple and incremental manner, as well as their easy and powerful application in reasoning processes (Valls, Gibert, Sánchez, & Batet, 2010). Other healthcare fields include decision support and information integration. Thus, the authors developed a medical treatment consent ontology as a preliminary notation to formalize the predicates used in creating consent rules.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Structure of the hierarchy overrides privilege to override the privacy disclosure decision



The rules used to model treatment consents are of the form "predicate$_1$(x_{11},x_{12},..,x_{1n}),… …., predicate$_m$(x_{m1},x_{m2},,,x_{mk}) → predicate(y_1,…y_p)" where the predicate instances stated left of the arrow symbol → are taken as the body of the rule and read as a conjunction of predicate$_1$($_{x11}$,x_{12},,,x_{1n})/\,…,/\ predicate$_i$(x_{i1},x_{i2},,,x_{ij})/\,…., /\predicate$_m$(x_{m1},x_{m2},,,x_{mk}) and the consequent is predicate(y_1,…,y_p). The intended meaning is that if the conditions specified in the antecedent hold, then the conditions specified in the consequent also hold. The authors show patient maturity evaluation rules for each state based on its consent laws as a case study in authors' published paper (Yu, Wijesekera, & Costa, 2014). In (Yu, Wijesekera, & Costa, 2014), the authors used open source software for implementing authors' framework YAWL Management System as the workflow management system and OpenMRS as authors' EMR system, Ontology Web Language (OWL) to describe authors' consent ontology, implemented using Protégé and applied Pellet as tour ontology reasoner. Similar techniques are used to obtain valid genetic privacy consents in addition to the privacy disclosure decision made by the highest override privilege as shown in Figure 7. As a result, additional task, which applied checking disclosure decision privilege rules, is added to genetic privacy consent service workflow.

Implementation

The authors developed a functioning prototype that addresses many aspects to develop an automated and integrated informed genetic information consent system. The prototype collages the data gathered during interactions with the medical provider with the application laws, regulations and policies to address the privacy issues specific to genetic information. The first component of implementing the genetic

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

privacy enforcement gathers the required information from an EMR system. This foundation allows the rules to be seamlessly integrated with the operations of the medical facility and maximizes the use of existing data to reduce additional data entry. The authors extended an open-sourced EMR by adding a tab to launch the consent workflow management system for an EMR system user. The workflow gathers additional information regarding aspects of the tasks being performed and then executes the path based on the purpose for accessing the genetic information (Disclosure, research, or treatment).

Once all the required information has been gathered from either the EMR database or from the users by invoking workflow-based prompts, the workflow management engine calls consent management system whenever an informed genetic consent is required. When a corresponding consent service is triggered, it implements an ontology described below to reason if the requested access is permitted or if not a consent form is required to generate. The access evaluation is done at each level in the hierarchy that is applicable to state, federal genetic privacy laws, and organization policies. By definition, federal laws are at the top of the hierarchy followed by state laws, local laws and then organizational policies. The hierarchical structure dictates how conflicts between laws and policies can be resolved based on the decisions made at each level. (For example, Federal Laws set a number of restrictions on the use of genetic data in research. Unless explicitly allowed in the Law, an organization cannot write conflicting policies.) The workflow engine evaluates the results from each level to return an overall decision as to whether access is allowed along with incorporating any other requirements that must be enforced (such as retention periods, written consent, de-identification, etc.)

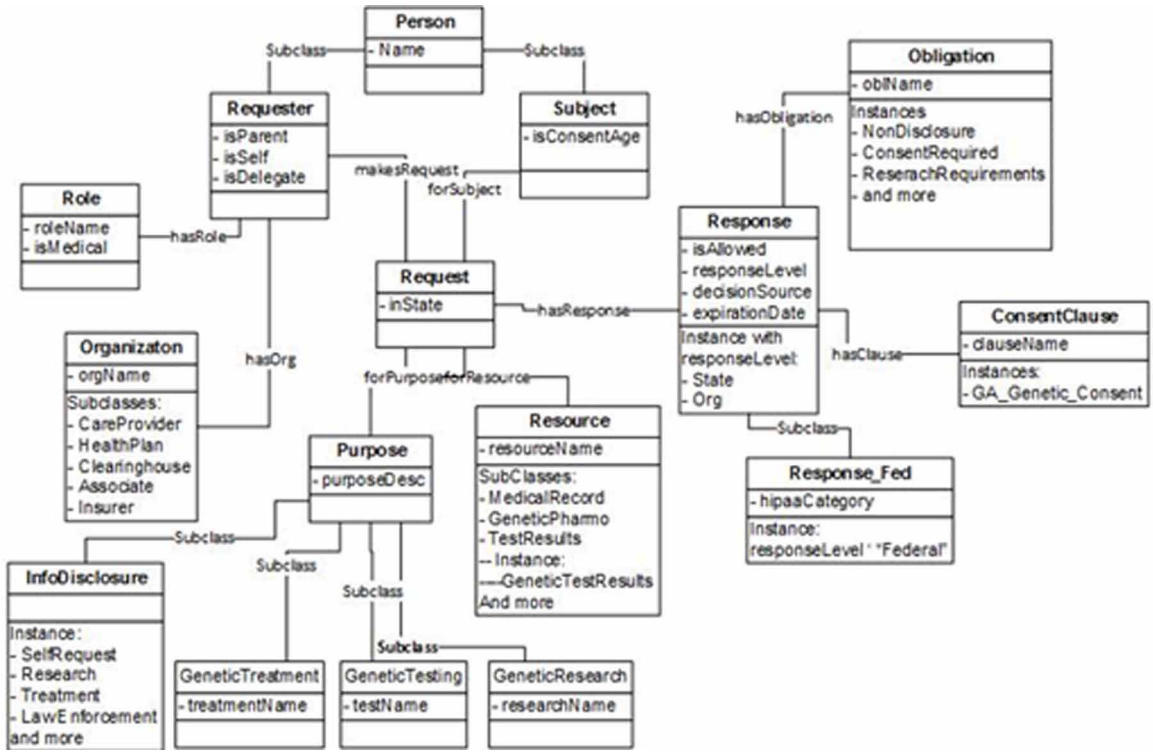
In order to support the consent service, the authors developed an ontology to capture the various aspects of enforcing privacy laws and policies. As shown in Figure 8, the ontology has four major entities:

- **Requester:** The person making the request to access the medical information including their role, associations with a specific organization, and information about this organization,
- **Resource:** The part of the electronic medical record being requested along with information about the subject (or patient)
- **Request:** Details on the purpose for requesting the information. The four specific purposes applicable to genetic information are disclosure, research, testing and treatment. The prototype currently implements the information disclosure component with the applicable specific instances for Self-Request by the Patient, Law Enforcement, etc.
- **Response:** The results of the reasoner applying the appropriate rules along with a list of any obligations that must be enforced by the EMR and specific consent clauses that are needed for the associated approvals. (A subclass for Federal Responses allows information about HIPAA-specific requirements to be gathered.)

The ontology diagram shown in Figure 8 contains some main properties and instances used in authors' work so far. The ontology is implemented using the Protégé platform with the laws (Federal and State), and the organizational policies are enforced using DLQuery and the Pellet reasoner. The predicate used in each rule uses the Request instance with the associated properties to gather additional information on the Requester and the Resource. The appropriate Response instance (Federal, State or Organization) receives the outcome of the rule regarding whether access is permitted or denied the specific law or policy that generated the result, and the appropriate obligations and clauses.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. The ontology diagram



The Override property implements the hierarchy of Federal/State laws and organizational policies described above. The override property is set as a meat-rule implementing a more strict law being evaluated that permits a looser law or policy to provide a different answer or supplemental information. So, if a Federal law permits access but has a clause where a state or organization can set a different policy to deny access, the override property in the Federal response is set to "True." Alternatively, the organizational policy might still permit access but add more clauses to the consent agreement that must be signed.

For example, HIPAA permits access to medical records even genetic data for treatment. In Georgia, there are additional obligations and consent requirements when the resource being accessed is from genetic testing. The Response structure allows both sets of answers to be passed back to the workflow engine for further evaluation and execution.

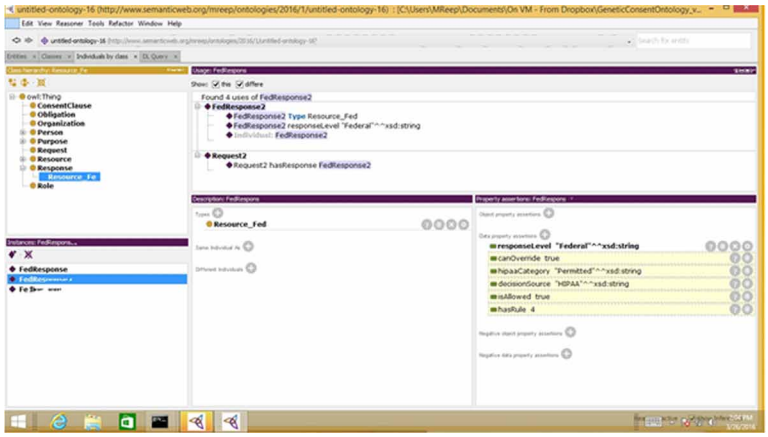
As an example of the implementation of obligations and overrides, the first DLQuery checks for the Federal law for access to the genetic Test Results for Treatment. The response is set to provide the corresponding information for the associated HIPAA category for Permitted, and therefore other laws can override the decision.

makesRequest(?r, ?req), forPurpose(?req, ?pur), purposeDesc(?pur, "Treatment"), hasResponse(?req, ?res), responseLevel(?res, "Federal") -> isAllowed(?res, true), canOverride(?res, true), hipaaCategory(?res, "Permitted"), decisionSource(?res, "HIPAA"), hasRule(?res, 4)

The next part of the example below shows the DL Query for the related request which sets the location in Georgia and that the rule can be executed if the Federal response allows an Override. The State response then allows access but with an additional obligation for a Consent Agreement and the agreement must have text specific to Georgia.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

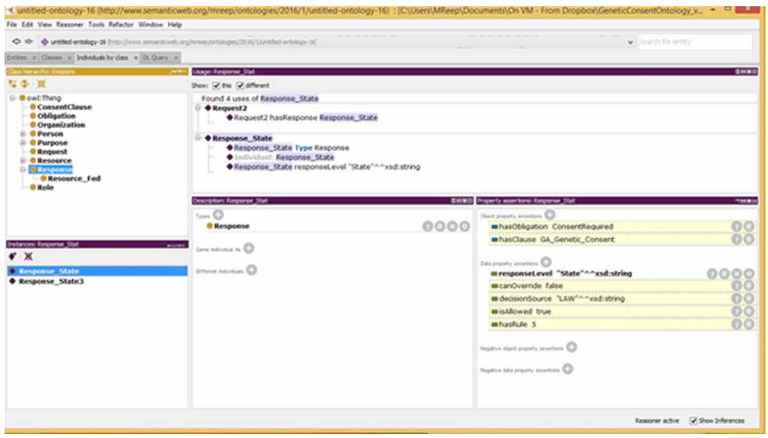
Figure 9. Outcome of the decision of accessing genetic data using Pellet reasoner



isSelf(?r, false), makesRequest(?r, ?req), inState(?req, “GA”), forResource(?req, ?resource), forPurpose(?req, ?pur), purposeDesc(?pur, “Treatment”), resourceName(?resource, “GeneticTestResults”), hasResponse(?req, ?res), responseLevel(?res, “Federal”), canOverride(?res, true), hasResponse(?req, ?resst), responseLevel(?resst, “State”), oblName(?obl, “ConsentRequired”), clauseName(?clause, “GAGeneticConsent”) -> isAllowed(?resst, true), canOverride(?resst, false), decisionSource(?resst, “GA_LAW”), hasObligation(?resst, ?obl), hasClause(?resst, ?clause)

Once the reasoner has evaluated the rules, the workflow engine will retrieve the Response, Obligation and Clauses instances through Java API calls to Protégé in order to enforce the results. The first part of the enforcement determines if access to the requested resource is allowed or denied. In order to consolidate the responses into one answer, first the Federal response is checked and, if an override is permitted, the state level is evaluated followed by the organizational level. If the consolidated response is determined to be a denial, then the EMR workflow ends with a message and information on the law or policy that generated the denial is provided. If the consolidated response is to allow the requested

Figure 10. Outcome of the decision of accessing genetic data using Pellet reasoner



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

access, the workflow displays information on the required obligations and requests acceptance to these conditions for data access. If consent is required, the applicable genetic privacy consent form will be retrieved for granting consent by legal consent giver. The workflow ends with access decision being provided to the requested genetic data.

Finally, the authors describe security issues, which are important considerations for any EMRs:

- **Access Control:** The medical team as a whole provides the required services to a patient who visits the medical center, from acceptance of a patient to the end of the treatment at the facility. Each team member plays a designated role in providing care with a set of assigned duties that are choreographed with each other, forming workflows. The team together provides the care planned for the patient. The authors used a role-based access control model to provide confidentiality.

Furthermore, enforced informed treatment consent and privacy consent can be viewed as more complex access controls with more complex rules.

- **Accountability:** To monitor the quality of care and be consistent with continuous improvements, an EMR system must enable auditing capabilities. HIPAA requires logging all PHI requests and the system actions for auditing. Authors' workflow-enforced EMR system with consent management allows the quality care team to review both procedures and outcomes from workflow logs and EMR system logs and consent logs, which provide an audit trail that satisfies accountability requirements.

CONCLUSION

Adding an executable consent-enforced service into workflow-based EMR systems to manage informed consent at runtime includes not only all benefits provided by workflow enforced EMR systems (Yu, & Wijesekera, 2013), but also provides significant benefits to medical practitioners and patients by reducing medical malpractice and potential medical treatment errors caused by missing informed consents (Yu, Wijesekera, & Costa, 2014). This work also improves the patient-caregiver relationship, risk management, decreases malpractice costs, enhances audit capabilities and fosters continuous care process improvement. Another advantage of authors' approach is the protection of patients' privacy by authors' EMR system. Additionally, this system can accommodate changes to policies and standards.

Solving the new consent management challenges in the area of genetic service, authors' approach allowed a single consent management system to handle different consent services based on the procedures' purposes. Authors' approach stores all genetic consents in the EMR system. Gaining consents and dealing with exceptions are recorded in the workflow management system. Thus, these records are readily available for quality of care audits and reviews.

Finally, using ontological techniques in a consent management system and creating a consent ontology enable keeping current of each State's evolving consent laws. Failure to track consent law changes closely and in near-real time is a sure recipe for failure.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Authors' work introduces an inspiring framework to extend existing EMR systems to avoid treatment errors caused by deviation from standardized treatment workflow and potential breach of privacy, or "psychosocial" harms to the individuals, as well, the many serious negatives that flow from lack of proper, informed consent.

Authors' ongoing work will expand this EMR system to automatically generate required consent rules from revised or newly enacted laws so that the operational EMR quickly reflect these changes, to allow patients to revoke consents and to appoint or re-appoint consent holders. This research path also includes further developing the evaluation rule repository to cover other more specific treatment scenarios.

REFERENCES

Faresi, A. L. A., Wijesekera, D., & Moidu, K. (2010). A comprehensive privacy-aware authorization framework founded on HIPAA privacy rules. In *Proceedings of the 1st ACM International Health Informatics Symposium.* ACM.

American Health Information Management Association. (2013). *Analysis of Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act; Other Modifications to the HIPAA Rules*. Retrieved August 16, 2014, from http://library.ahima.org/PdfView?oid=106127

American Society of Human Genetics Social Issues Subcommittee on Familial Disclosure. (1998). ASHG STATEMENT Professional Disclosure of Familial Genetic Information. *American Journal of Human Genetics*, *62*(2), 474–483. doi:10.1086/301707 PMID:9537923

Anonymous. (2007). *Healthcare's most expensive piece of paper*. DialogMEDICAL. Retrieved May 10, 2013, from http://www.dialogmedical.com/wp- content/uploads/2010/04/ROI%20White%20Paper%2005-07.pdf

Appelbaum, P. S., & Thomas, G. (1995). The MacArthur Treatment Competence Study. I: Mental illness and competence to consent to treatment. *Law and Human Behavior*, *19*(2), 105–126. doi:10.1007/BF01499321 PMID:11660290

Ayuso, C., Millán, J. M., Mancheño, M., & Dal-Ré, R. (2013). Informed consent for whole-genome sequencing studies in the clinical setting. Proposed recommendations on essential content and process. *European Journal of Human Genetics*, *21*(10), 1054–1059. doi:10.1038/ejhg.2012.297 PMID:23321621

Beauchamp, T. L., & Childress, J. F. (1989). *Principles of Biomedical Ethics* (3rd ed.). New York: Oxford University Press.

Belmont, J., & McGuire, A. L. (2009). The futility of genomic counseling: Essential role of electronic health records. *Genome Medicine*, *1*(5), 1. doi:10.1186/gm48 PMID:19439060

Bergmann, J., Bott, O. J., Pretschner, D. P., & Haux, R. (2007). An econsent-based shared EHR system architecture for integrated healthcare networks. *International Journal of Medical Informatics*, *76*(2-3), 130–136. doi:10.1016/j.ijmedinf.2006.07.013 PMID:16971171

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bhatti, R., & Grandison, T. (2007). *Towards improved privacy policy coverage in healthcare using policy refinement*. Berlin: Springer. doi:10.1007/978-3-540-75248-6_11

Boonstra, H., & Nash, E. (2000). Minors and the right to consent to health care. *The Guttmacher Report on Public Policy*, *3*(4), 4–8. PMID:12134879

Bourgeois, F. C., Taylor, P. L., Emans, S. J., Nigrin, D. J., & Mandl, K. D. (2008). Whose personal control? Creating private, personally controlled health records for pediatric and adolescent patients. *Journal of the American Medical Informatics Association*, *15*(6), 737–743.

Buchanan, A. (2004). Mental capacity, legal competence and consent to treatment. *Journal of the Royal Society of Medicine*, *97*(9), 415–420.

Clayton, E. W., Steinberg, K. K., Khoury, M. J., Thomson, E., Andrews, L., Kahn, M. J. E., ... Weiss, J. O. (1995). Informed consent for genetic research on stored tissue samples. *Journal of the American Medical Association*, *274*(22), 1786–1792. doi:10.1001/jama.1995.03530220052033 PMID:7500511

Coiera, E., & Clark, R. (2004, March/April). e-Consent: The design and Implementation of consumer consent mechanisms in an electronic environment. *Journal of the American Informatics Association*, *11*(2). PMID:14662803

Compbell, D. (June, 2004). Case and Commentary – Listen to the family. *Morbidity & Mortality Rounds on the Web. AHRQ.* Retrieved August 8, 2013, from http://webmm.ahrq.gov/case.aspx?caseID=62

Dang, J., Hedayati, A., Hampel, K., & Toklu, C. (2008). An ontological knowledge framework for adaptive medical workflow. *Journal of Biomedical Informatics*, *41*(5), 829–836. doi:10.1016/j.jbi.2008.05.012 PMID:18602872

Del Carmen, M. G., & Joffe, S. (2005). Informed consent for medical treatment and research: A review. *The Oncologist*, *10*(8), 636–641. doi:10.1634/theoncologist.10-8-636 PMID:16177288

Health insurance portability and accountability act. U.S. department of health and human services. (2013). Retrieved February 20, 2014, from http://www.gpo.gov/fdsys/pkg/FR-2013-01-25/pdf/2013-01073.pdf

Fisher, C. B. (2002). Goodness-of-Fit Ethic for Informed Consent. *A. Fordham Urb. LJ*, *30*, 159. PMID:15868667

Garde, S., Knaup, P., Hovenga, E. J., & Heard, S. (2007). Towards Semantic Interoperability for Electronic Health Records--Domain Knowledge Governance for open EHR Archetypes. *Methods of Information in Medicine*, *46*(3), 332–343. PMID:17492120

Genetic Information Non-discrimination Act of 2008 (GINA). *Pub. L. 110-233*, 122 Stat. 883, codified as amended in scattered sections of 26, 29, and 42 U.S.C.

Gostin, L. O., & Hodge, J. G. Jr. (1999). Genetic privacy and the law: An end to genetics exceptionalism. *Jurimetrics*, 21–58. PMID:12465649

Grimson, J., Stephens, G., Jung, B., Grimson, W., Berry, D., & Pardon, S. (2001). Sharing health- care records over the internet. *IEEE Internet Computing*, *5*(3), 49–58. doi:10.1109/4236.935177

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Grisso, T., & Appelbaum, P. S. (1998). *Assessing competence to consent to treatment: A guide for physicians and other health professionals*. Oxford University Press.

Hall, D. E., Prochazka, A. V., & Fink, A. S. (2012). Informed consent for clinical treatment. *Canadian Medical Association Journal*, *184*(5), 533–540. doi:10.1503/cmaj.112120 PMID:22392947

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. doi:10.1006/knac.1993.1008

Hallinan, D., & Friedewald, M. (2015). Open consent, biobanking and data protection law: can open consent be 'informed'under the forthcoming data protection regulation?. *Life Sciences, Society and Policy, 11*(1), 1.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, *43*(5), 907–928. doi:10.1006/ijhc.1995.1081

Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science*, *339*(6117), 321–324. doi:10.1126cience.1229566 PMID:23329047

HIPAA Compliance Assistance. (2003). *Summary of the HIPAA privacy rule*. Office for Civil Rights. Retrieved February 20, 2014, from http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/privacysummary.pdf

Holder, A. R. (1987). Minors' rights to consent to medical care. *Journal of the American Medical Association*, *257*(24), 3400–3402. doi:10.1001/jama.1987.03390240106033 PMID:3586272

Kaye, J., Heeney, C., Hawkins, N., De Vries, J., & Boddington, P. (2009). Data sharing in genomics—re-shaping scientific practice. *Nature Reviews. Genetics*, *10*(5), 331–335. doi:10.1038/nrg2573 PMID:19308065

Kerath, S. M., Klein, G., Kern, M., Shapira, I., Witthuhn, J., Norohna, N., ... Taioli, E. (2013). Beliefs and attitudes towards participating in genetic research–a population based cross-sectional study. *BMC Public Health*, *13*(1), 1. doi:10.1186/1471-2458-13-114 PMID:23388465

Khoury, M. J., & Dorman, J. S. (1998). The human genome epidemiology network. *American Journal of Epidemiology*, *148*(1), 1–3. doi:10.1093/aje/148.1.1 PMID:9663396

Knoppers, B. M. (2005). Consent revisited: Points to consider. *Health Law Review*, *13*(2-3), 33–38. PMID:16459412

Liao, S. M. (2009). Is there a duty to share genetic information? *Journal of Medical Ethics*, *35*(5), 306–309. doi:10.1136/jme.2008.027029 PMID:19407036

Locatelli, P., Restifo, N., Gastaldi, L., & Corso, M. (2012). Health Care Information Systems: Architectural Models and Governance. In C. Kalloniatis (Ed.), *Innovative Information Systems Modelling and Techniques* (pp. 71–96). Rijeka, Croatia: InTech. doi:10.5772/38212

Lunshof, J. E., Chadwick, R., Vorhaus, D. B., & Church, G. M. (2008). From genetic privacy to open consent. *Nature Reviews. Genetics*, *9*(5), 406–411. doi:10.1038/nrg2360 PMID:18379574

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mascalzoni, D., Hicks, A., Pramstaller, P., & Wjst, M. (2008). Informed consent in the genomics era. *PLoS Medicine*, *5*(9), e192. doi:10.1371/journal.pmed.0050192 PMID:18798689

Matos, E. E., Campos, F., Braga, R., & Palazzi, D. (2010). CelOWS: An ontology based framework for the provision of semantic web services related to biological models. *Journal of Biomedical Informatics*, *43*(1), 125–136. doi:10.1016/j.jbi.2009.08.008 PMID:19695346

Merker, J. (2014). *Establishing a clinical genomics program at an academic medical center.* Regional APC/PDAS Meeting, Kauuai.

McGuire, A. L., & Beskow, L. M. (2010). Informed consent in genomics and genetic research. *Annual Review of Genomics and Human Genetics*, *11*(1), 361–381. doi:10.1146/annurev-genom-082509-141711 PMID:20477535

Mulholland, W. F. (1998). Genetic privacy and discrimination: A survey of state legislation. *Jurimetrics*, *39*, 317. PMID:11762444

National Conference of State Legislatures. (2008). *Genetic privacy laws.* Congressional Research Service. Retrieved March 2008 from http://www.ncsl.org/research/health/genetic-privacy-laws.aspx

Németh, A. H., Kwasniewska, A. C., Lise, S., Schnekenberg, R. P., Becker, E. B., Bera, K. D., ... Talbot, K. (2013). Next generation sequencing for molecular diagnosis of neurological disorders using ataxias as a model. *Brain*, awt236. PMID:24030952

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Retrieved February 20, 2014, from http://protege.stanford.edu/publications/ontology_development/ontology101.pdf

O'Daniel, J., Scott, J., Varga, E., Speights, V., & Gordon, E. (2012). Lecture iv: Genomic medicine: Communicating with the patient. *TRiG Curriculum: Lecture 4.*

O'Keefe, C. M., Greenfield, P., & Goodchild, A. (2005). A decentralized approach to electronic consent and health information access control. *Journal of Research and Practice in Information Technology*, *37*(2), 161–178.

O'Leary, C. (2010). Informed consent-principles and practice. *ASA Newsletter*, *74*(2), 20–21.

Petersen, A. (2007). 'Biobanks'" engagements": Engendering trust or engineering consent?'. *Genomics. Social Policy*, *3*(1), 1–13.

Petronio, S. (2012). *Boundaries of privacy: Dialectics of disclosure*. Suny Press.

Presidential Commission for the Study of Bioethical Issues. (2013). Retrieved March 23, 2015, from http://www.bioethics.gov/sites/default/files/Informed%20Consent%20Background%20090413.pdf

Prince, A. E., & Roche, M. I. (2014). Genetic information, non-discrimination, and privacy protections in genetic counseling practice. *Journal of Genetic Counseling*, *23*(6), 891–902. doi:10.100710897-014-9743-2 PMID:25063358

Rehm, H. L., Hynes, E., & Funke, B. H. (2016). The changing landscape of molecular diagnostic testing: Implications for Academic Medical Centers. *Journal of Personalized Medicine, 6*(1), 8.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Riaño, D., Real, F., López-Vallverdú, J. A., Campana, F., Ercolani, S., Mecocci, P., & Caltagirone, C. (2012). An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. *Journal of Biomedical Informatics*, *45*(3), 429–446. doi:10.1016/j.jbi.2011.12.008 PMID:22269224

Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., & Kim, D. (2015). Methods of integrating data to uncover genotype-phenotype interactions. *Nature Reviews. Genetics*, *16*(2), 85–97. doi:10.1038/nrg3868 PMID:25582081

Roach, J. C., Glusman, G., Smit, A. F., Huff, C. D., Hubley, R., Shannon, P. T., ... Shendure, J. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, *328*(5978), 636–639. doi:10.1126cience.1186802 PMID:20220176

Rosse, C., & Mejino, J. L. Jr. (2003). A reference ontology for biomedical informatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*, *36*(6), 478–500. doi:10.1016/j.jbi.2003.11.007 PMID:14759820

Rothstein, M. A. (2014). Privacy and Confidentiality. Routledge Handbook of Medical Law and Ethics, 52.

Ruan, C., & Yeo, S. S. (2009). Modeling of an Intelligent e-Consent System in a Healthcare Domain. *J. UCS*, *15*(12), 2429–2444.

Russello, G., Dong, C., & Dulay, N. (2008). Consent-based workflows for healthcare management. In *Policies for Distributed Systems and Networks, POLICY. IEEE Workshop on* (pp. 153-161). 10.1109/POLICY.2008.22

Salgo v. Leland Stanford, Jr., Univ. Bd. of Trustees, 317 P. 2d 170, 181 (Cal. App. Ct. 1957).

Saranummi, N. (2005). PICNIC architecture. *Studies in Health Technology and Informatics*, *115*, 37–60.

Scheuner, M. T., de Vries, H., Kim, B., Meili, R. C., Olmstead, S. H., & Teleki, S. (2009). Are electronic health records ready for genomic medicine? *Genetics in Medicine*, *11*(7), 510–517. doi:10.1097/GIM.0b013e3181a53331 PMID:19478682

Secko, D. M., Preto, N., Niemeyer, S., & Burgess, M. M. (2009). Informed consent in biobank research: A deliberative approach to the debate. *Social Science & Medicine*, *68*(4), 781–789. doi:10.1016/j.socscimed.2008.11.020 PMID:19095337

Sheppard, N. P., Safavi-Naini, R., & Jafari, M. (2009). A digital rights management model for healthcare. In Policies for Distributed Systems and Networks. *POLICY 2009. IEEE International Symposium* (pp. 106-109). 10.1109/POLICY.2009.8

Sherlock, E. (2009). *Disclosure of patient's genetic information without their consent- Is the "public interset" really a Sufficient Justification?* Genomics Law Report. retrieved March 2, 2015, from http://www.genomicslawreport.com/index.php/2009/11/10/disclosure-of-patients-genetic-information-without-their-consent-is-the-public-interest-really-a-sufficient-justification/

Shortliffe, E. H., & Cimino, J. J. (2006). *Biomedical informatics*. Springer Science + Business Media, LLC.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Strouse, D. S. (2005). Informed consent to genetic research on banked human tissue. *Jurimetrics*, 135–152.

Struewingl, J. P., & Abeliovich, D. (1995). The carrier frequency of the BRCA1 185delAG mutation. *Nature Genetics*, 11.

Studdert, D. M., Mello, M. M., Levy, M. K., Gruen, R. L., Dunn, E. J., Orav, E. J., & Brennan, T. A. (2007). Geographic variation in informed consent law: Two standards for disclosure of treatment risks. *Journal of Empirical Legal Studies*, *4*(1), 103–124. doi:10.1111/j.1740-1461.2007.00083.x

Ullman-Cullere, M. H., & Mathew, J. P. (2011). Emerging landscape of genomics in the electronic health record for personalized medicine. *Human Mutation*, *32*(5), 512–516. doi:10.1002/humu.21456 PMID:21309042

Valls, A., Gibert, K., Sánchez, D., & Batet, M. (2010). Using ontologies for structuring organizational knowledge in Home Care assistance. *International Journal of Medical Informatics*, *79*(5), 370–387. doi:10.1016/j.ijmedinf.2010.01.012 PMID:20185360

VHA HANDBOOK 1004.05, Transmittal Sheet. (2005). *IMEDCONSENT*™. Retrieved August 17, 2013, from http://www.va.gov/vhapublications/ViewPublication.asp?pub_ID=1857

Virginia women's center - cancer family history questionnaire. (n.d.). Retrieved May 22, 2015, from https://www.virginiawomenscenter.com/assets/HEREDITARY_CANCER_RISK_ASSESSMENT_FORM.pdf

Wickramasinghe, N., & Geisler, E. (Eds.). (2008). *Encyclopedia of Healthcare Information Systems*. Hershey, PA: Medical Information Science Reference. doi:10.4018/978-1-59904-889-5

Wirth, T., Parker, N., & Ylä-Herttuala, S. (2013). History of gene therapy. *Gene*, *525*(2), 162–169. doi:10.1016/j.gene.2013.03.137 PMID:23618815

Win, K., Song, H., Croll, P., & Cooper, J. (2002). Implementing patients consent in electronic health record systems. *Proceedings of CollECTeR*.

Win, K. T., & Fulcher, J. A. (2007). Consent mechanisms for electronic health record systems: A simple yet unresolved issue. *Journal of Medical Systems*, *31*(2), 91–96. doi:10.100710916-006-9030-3 PMID:17489500

Yu, A. C. (2006). Methods in biomedical ontology. *Journal of Biomedical Informatics*, *39*(3), 252–266. doi:10.1016/j.jbi.2005.11.006 PMID:16387553

Yu, B., & Wijesekera, D. (2013). Building Dialysis Workflows into EMRs. *Procedia Technology, 9*, 985-995.

Yu, B., Wijesekera, D., & Costa, P. (2014). Consent-Based Workflow Control in EMRs. *HCIST- International Conference on Health and Social Care Information Systems and Technologies,* 72-79.

Yu, B., Wijesekera, D., & Costa, P. (2014). An Ontology for Medical Treatment Consent. *The 9th International Conference on. Semantic Technologies for Intelligence, Defense, and Security - STIDS 2014*.

Yu, B., Wijesekera, D., Costa, P., & Hiremagalore, S. (2015). Genetic Counseling Using Workflow-based EMRs. *The 10th International Conference on. Semantic Technologies for Intelligence, Defense, and Security - STIDS 2015*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Consent Ontology: An ontology of consent.
Electronic Medical Record (EMR) System: An electronic system that manages medical data.
Genetic Counseling: Counseling related to genetic issues.
Genetic Services: A set of services that generates or uses genetic data.
Medical Informed Consent: A decision and permission to perform a procedure in medicine.
Participating Research Consent: A decision and permission to participate in medical research.
Privacy Consent: A decision and permission to use and disclose information.
The Health Insurance Portability and Accountability Act (HIPAA): The federal laws to protect patients' PHI for covered entities.
Treatment Consent: A decision and permission to treat medically.
Workflow Management: A management system for executing workflows.

This research was previously published in Design, Development, and Integration of Reliable Electronic Healthcare Platforms edited by Anastasius Moumtzoglou; pages 211-242; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 62
Towards User Authentication Requirements for Mobile Computing

Yaira K. Rivera Sánchez
University of Connecticut, USA

Steven A. Demurjian
University of Connecticut, USA

ABSTRACT

The emergence and ubiquity of mobile computing has placed powerful capabilities in one's hand providing a wide range of applications such as email, calendar, photos, browsers, social network, communication, shopping, health and fitness, games etc., which were once restricted to traditional platforms. Such applications on a single mobile device raise critical security issues related to managing identity, re-authenticating users that stay active for long periods of time, protecting sensitive PII and PHI against access and misuse, insuring secure transactions, and protecting the physical device. This chapter explores user authentication requirements for mobile computing by: evaluating alternative user authentication requirements in order to make recommendations on their usage in authentication; identifying authentication methods used in mobile healthcare applications; and proposing a set of requirements for user authentication to handle the situation when a user seeks to be securely authenticated across a set of applications that are placed into context within a framework.

INTRODUCTION

Mobile computing platforms have greatly increased in the marketplace. By the end of 2013, globally, 6% of the population owned a tablet and 22% owned a smartphone (Jeggenstuen, 2013). In the United States, by the end of 2014, there was a 63.5% penetration rate (eMarketer, 2013). Such growth leads to the reduction of traditional desktop and laptop sales (Gartner, 2014). As individuals rapidly change the ways they access data and information, there is a profound impact on the way that users must be

DOI: 10.4018/978-1-7998-1204-3.ch062

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

authorized and authenticated. Accessing data and executing apps on a mobile platform is substantially more dynamic than traditional computing on laptops and desktops. A mobile device user may have a myriad of open applications including email accounts (corporate, Gmail, yahoo, etc.), browsers, social network apps (Facebook, Twitter, LinkedIn, etc.), communication apps (Skype, Snapchat, SMS, etc.), shopping apps (Amazon, JCPenney, Walgreens, etc.), health and fitness apps (CVS Health, Microsoft HealthVault, MyQuest, etc.), games, etc. In such a setting, users are often logged onto and authenticated to all of these apps simultaneously and move among them throughout the day, providing them with the ability to store, exchange, and view both personally identifiable information (PII) and personal health information (PHI). The drawback is the need for users to be authenticated on each individual application and often re-authenticated during sessions that can last for long periods of time. The challenge in this situation is to deliver a means for user authentication for mobile computing in order to provide a more seamless experience for users.

As a result, for a user to manage his/her personal information in several domains, we need to find a secure means while maintaining their privacy. Therefore, the challenge is a need to provide a cohesive means of user authentication that can cross all of these different domain boundaries with high availability of data, while still providing the necessary level of security and privacy. Our approach to such a challenge is to identify existing frameworks, standards, types of access control, and authentication methods that can identify requirements for user authentication for mobile devices and applications that would then be able to interact and share information with other mobile, cloud, and computing applications for a given domain. In the case of healthcare, health information technology (HIT) systems have emerged (e.g., electronic medical records (EMRs), practice management systems (PMSs) and personal health records (PHRs)), due to the American Recovery and Reinvestment Act (ARRA) of 2009 (U.S. Department of Education, 2010), all which must adhere to the Health Insurance Portability and Accountability Act (HIPAA) (HHS.gov, 2013) for the security, availability, transmission, and release of a patient's medical information. Note that an EMR is the primary means for making patient information electronically available to medical providers for healthcare. This wide adoption of HIT allows individuals to track their own health data independent of their provider. For example, the fitness market has exploded with wearable fitness devices that link to mobile applications, e.g., Apple has a new HealthKit app (Apple iOS 9, 2014) that utilizes a dashboard to manage health and fitness data while Google has the Fit fitness tracker (Real IT Experts, 2013). Both have pushed strongly into the smartwatch market to track activity, heart rate, blood pressure, etc. (Kelly, 2014) and these and other apps must be able to securely exchange HIPAA protected information back and forth with mobile applications for both patients and medical providers. In such a scenario, the same information is stored in different ways in diverse HIT systems and other applications, and this information must be staged back and forth to a mobile device and be protected from theft and misuse while simultaneously allowing a user to share the data with authorized medical providers and others, such as family members.

In support of this, the main focus of this work is to identify the requirements and capabilities that are needed to support user authentication that allows a person utilizing a mobile device to gain access to protected information from multiple input sources (cloud services, other users, other apps, etc.), with the health care domain as a test-bed. This chapter identifies and explains user authentication requirements for mobile computing that offers a high degree of security assurance and that is able to facilitate the aforementioned discussion on the access, usage, and sharing of information among users. In order to support user authentication in mobile computing there are key security concepts to be considered. In terms of *credentials*, mobile applications have long used passwords as a main means of access, with recent OS

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

mobile operating system updates providing support for fingerprints and the proposal of single-sign on in a mobile setting (Yu, Wang, & Mu, 2012). *Risk-based authentication* (Abe, Itoh, & Takahashi, 2007) offers the potential to assess the risk via a geographic location log, the app being accessed, the time of day a user typically logs into the app, a device's IP address, etc.; such geographic information may be useful when users stay logged on for long periods of time. With respect to *biometrics*, one approach has proposed the tracking of graphical touches on a mobile device's screen (pressure, intensity, and touch) as a means to authenticate a user's identity (Zhao, Feng, & Shi, 2013). Other techniques include: *multi-factor authentication* (Bhargav-Spantzel et al., 2007), which would allow a mobile device to authenticate using something that a user knows (e.g., username/password, PIN, image, etc.), possesses (e.g., token or certificate), or is (e.g., fingerprint, retina scan, etc.); *user profiling* (Jeong & Choi, 2012), which discovers patterns in large quantities to establish if the user is who he/she says he/she is by tracking the way that he/she utilizes his/her mobile device, and deviations of behavior may indicate a problem; and, utilizing keystroke in conjunction with handwriting-based analysis (also multi-factor) to achieve a higher degree of confidence in the authentication process and user identity (Trojahn & Ortmeier, 2013).

The chapter explores user authentication requirements for mobile computing, accompanied by a case study of its usage for the aforementioned health care scenario by: investigating and presenting work by researchers in the area; by identifying what has been done in security for authentication, in general, and for the mobile case, to identify gaps of work that exist in trying to streamline the authentication process; by examining various mobile applications in the healthcare domain to understand their authentication approaches; by identifying and discussing limitations of the varied authentication approaches with regards to the sharing and exchange of information; and, by analyzing the approaches in terms of what they utilize in order to do user authentication. The end result of the chapter is an explanation and identification of user authentication requirements for mobile devices and applications that must securely interact with numerous devices and systems to support security assurance for a domain, which in our case will be demonstrated in healthcare.

The remainder of this chapter has 6 sections. In the *Background and Motivation* section, we state why healthcare has been chosen as a test bed, and review HIT systems and applications including Electronic Medical Records (EMRs) and Personal Health Records (PHRs) along with a scenario on futuristic needs in terms of secure access to health data. The *Security Concepts Impacting Mobile Authentication* section explores the security concepts needed to understand the different components of user authentication, focusing on three main authentication factors: knowledge factors, possession factors, and inherence factors. The *Authentication in Mobile Healthcare Applications* section presents several existing mobile healthcare applications, focusing on their respective methods of authentication to provide access to a user. The *Authentication Requirements in Mobile Computing* section identifies and discusses a set of five requirements: long-term mobile application sessions; beyond two-factor authentication; advanced threats and hacking; multiple and simultaneous authentication; and, lost or stolen device. These five requirements are then placed into context through the presentation of a framework that demonstrates their interactions. Then, the *Future Trends* section identifies approaches that are emerging and may impact authentication, such as: the increase of multi-factor authentication in mobile devices (from both Google and Apple) (Beranek, 2014), the concept of continuous authentication (Schaffer, 2015), new approaches for mobile security (like VIP Access (Symantec, 2013)), and new mobile operating systems for mobile devices (e.g., Firefox OS, Tizen, and Ubuntu) (Boxall, 2013). Finally, the *Conclusion* section finishes this chapter.

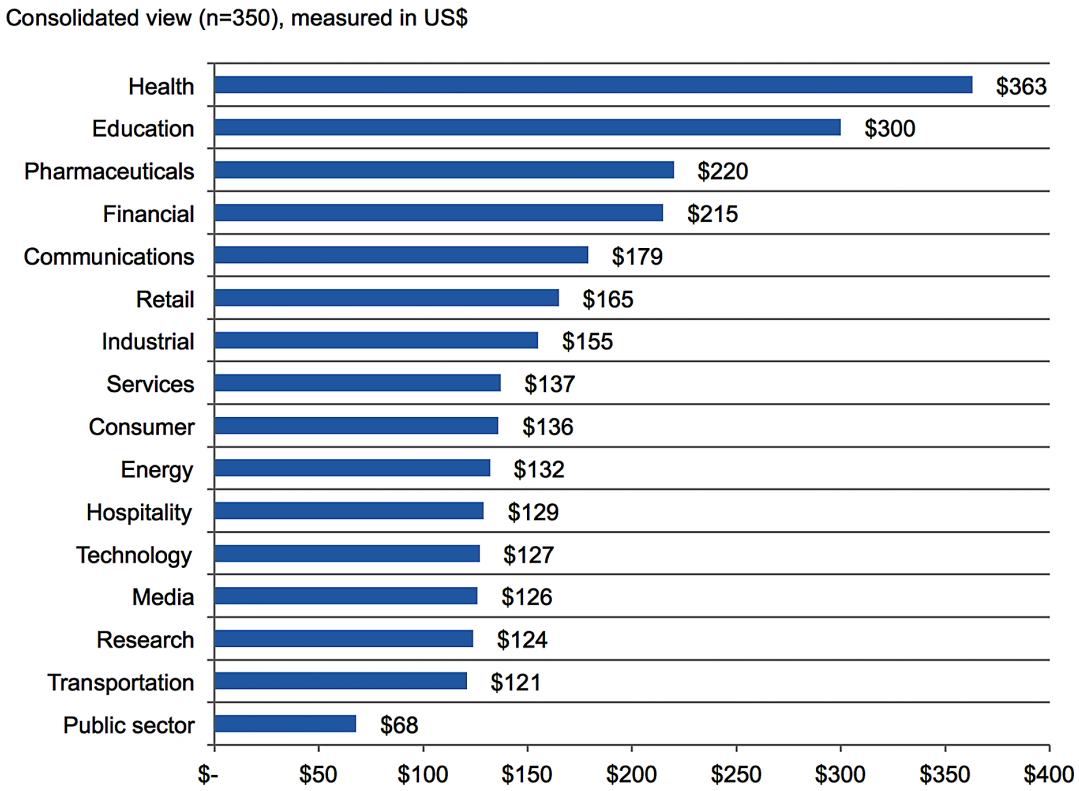
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND AND MOTIVATION

As previously stated, we will focus on identifying and explaining user authentication requirements utilizing a healthcare setting as base. We chose the healthcare domain since improper disclosure of data can have serious impact on patients, which can include personal embarrassment, prejudice, ostracization from family and community groups, death, as well as issues with insurability (Rindfleisch, 1997). Moreover, a report made by Ponemon Institute (2009) revealed that the cost of data breaches in the majority of the industries (e.g. communications, retail) per lost or stolen record averages approximately $154 per record (see Figure 1) (Ponemon Institute, 2015) . Nevertheless, in the healthcare industry, the cost per lost or stolen record could be as high as $363 per record, which is more than double the cost to manage a data breach in other industries. We consider that utilizing a healthcare setting as basis is an appropriate and informative manner to present the main focus of the chapter; note that the proposed requirements can be applied to other domains as well.

To start the discussion, we define various concepts that will appear frequently throughout the chapter and present a case scenario to have an idea of how security can be applied to these concepts. We divide the presentation into two parts. First, we review health information technology (HIT) systems utilized by patients, medical providers, hospitals, etc. Then, we provide a plausible, and at times, futuristic scenario for the usage of mobile devices and their applications, in general, and as related to health care.

Figure 1. Per capita cost by industry classification.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Health IT Systems Concepts

To highlight the need for authentication, consider the usage and management of personal health information in the healthcare domain. In the last several years, various HIT systems have emerged such as: electronic health records (EHRs), practice management systems (PMSs) and personal health records (PHRs) (Rouse, 2015). Such automated and interoperable systems are utilized to help reduce errors and lower costs, in comparison to paper-based systems. An EHR contains health information about an individual patient, which can be shared throughout different healthcare settings (Stricker, 2010). Such data can include the demographics, medical history, and allergies of the patient. A PMS can be defined as applications and tools utilized in health care institutions in order to keep track of all the operations these do such as patient billing and medical data processing (Codjia, 2010). Finally, we have electronic medical records (EMRs), which are mainly utilized by organizations that deliver care (e.g. hospitals and clinics) to manage records. Usually, the terms EHR and EMR are utilized interchangeably.

The emergence of the systems described above and of other HIT systems is mostly due to the American Recovery and Reinvestment Act (ARRA) of 2009 (U.S. Department of Education, 2010) to assist the United States economy to recuperate from an economic recession by allocating $787 billion to different areas, with $150 billion to healthcare industry, of which $36 billion was allocated to states to focus on doing HIT adoption including EMRs, PHRs, and Health Information Exchange (HIE). All of the systems must adhere to the Health Insurance Portability and Accountability Act (HIPAA) for the security, availability, transmission, and release of a patient's medical information. Also, in order for practitioners to obtain incentives for implementing EHRs in their medical offices, they must follow several requirements. These are described under the Meaningful Use stage 1 and stage 2 criteria (CMS.gov, 2015). Meaningful Use Criteria are utilized to certify the capabilities of EHRs in order to improve the quality, safety, and efficiency of care, to engage patients in the care process, and to maintain privacy and security of patient health data. EHRs and EMRs in the United States are required to meet Stage 2; Stage 3 objectives were recently released by Centers for Medicare and Medicaid (CMS) (Centers for Medicare & Medicaid Services, 2015). This wide adoption of HIT allows individuals to track their own health data independent of their provider. In addition to the Health IT systems described above, there exist many mobile applications that can be utilized to track a patient's well being, to manage their health data, among many other features. Also, the fitness market has exploded with wearable devices that link to mobile applications, e.g., Apple has a new HealthKit app (Apple iOS 9, 2014) that utilizes a dashboard to manage health and fitness data while Google has the Fit fitness tracker (Real IT Experts, 2013). Both have pushed strongly into the smartwatch market to track activity, heart rate, blood pressure, etc. (Kelly, 2014). As more applications store healthcare data in different systems, it becomes difficult for users to obtain their data seamlessly. As a result, we are faced with numerous issues for user authentication for mobile applications that must be addressed for the user to access his/her data in a more secure and seamless way.

Futuristic Case Scenario

Consider that accessing data and executing apps on a mobile platform is substantially more dynamic than traditional computing on laptops and desktops. A mobile device user may have a myriad of open applications including email accounts (corporate, Gmail, Yahoo Mail, Hotmail, etc.), browsers (Chrome, Safari, Opera, Firefox, etc.), financial management apps (banking, insurance, investment, etc.), social

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

network apps (Facebook, Twitter, Instagram, LinkedIn), communication apps (Skype, Snapchat, SMS, etc.), shopping apps (Amazon, Barnes & Noble, JCPenney, Home Depot, Walmart, etc.), loyalty programs to encourage product usage (Starbucks, Dunkin Donuts, eBates, etc.), health and fitness apps (CVS Health, Microsoft HealthVault, MyQuest, Apple Healthkit, Google Fit, Fitbit, etc.), video-streaming (YouTube, Netflix, HBO GO, Hulu, Vudu, etc.), games, etc. In such a setting, users are often logged onto and regularly stay authenticated to all of these apps simultaneously and move among them throughout the day keeping sessions open (logons active) for long periods of time, providing them with the ability to store, exchange, and view information that ranges from the least confidential (video streaming) to sensitive personal information (PII and PHI). The drawback is the need for users to be authenticated on each individual application and often re-authenticated during sessions that can last for long periods of time. Therefore, one of the challenges in this scenario is to provide a means for identifying and explaining user authentication requirements for mobile computing in order to provide a more seamless experience for users.

Suppose we have a set of IT applications, systems, and databases for a particular domain, and there will be mobile applications for different classes of users that are interested in both reading/writing data in one or more of these IT Apps, Systems, DBs (ITASDB). Each of the ITASDB has some sort of public API or cloud services that facilitate access and usage. For example in the case of health care, there would be cloud computing services to facilitate access to EHRs, PHRs, pharmacy systems, imaging and testing laboratory results, etc. What unique type of infrastructure would be necessary to allow a mobile application to easily interact with one or more ITASDB via cloud services where what they interact with can shrink and grow over time? What do we have to do in order to allow mobile apps (which may be the user or may be the app itself) to register in such a way as to get access to ITASDB initially and then change over time? And once registered, can they easily get access to cloud services in order to perform the appropriate and relevant read/write operations? In a health care scenario, it would be reasonable in the very near future for a person to manage medical and fitness information to: get information from HIT systems, fitness apps, etc., for a complete integrated view of health and personal data; and, share information from the mobile applications to various stakeholders such as family members (elder care), personal trainers or nutritionists, and medical providers.

SECURITY CONCEPTS IMPACTING MOBILE AUTHENTICATION

This section reviews several key security concepts to allow the reader to understand their usage and impact in regards to authentication. We focus on *user authentication*, which is the process that is performed to determine if a user is whom he/she says. There are three main authentication factors that we emphasize, which can be utilized to verify a user's identity: knowledge factors, possession factors, and inherence factors. Table 1 summarizes the factors in terms of: definition, examples, and advantages/disadvantages. In addition, Table 1 provides information on two-factor/multi-factor authentication, which consists on the combination of two or more of the previously mentioned authentication factors.

While there has been an increase in usage of biometrics and multi-factor authentication, there has also been research and development endeavors that have identified the potential pitfalls and have proposed the inclusion of user profiling and risk-based authentication for mobile security frameworks. Recall that *user profiling* consists of discovering patterns or correlations in large quantities to establish if the user is who he/she says he/she is, where the way that a user utilizes his/her mobile device would

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Main Authentication factors.

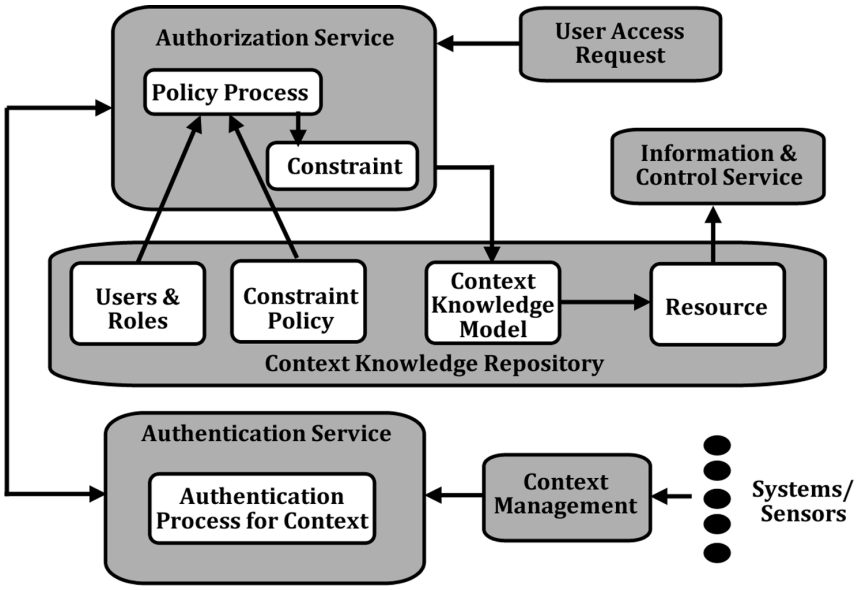
	Knowledge Factors	Possession Factors	Inherence Factors	Two-Factor/ Multi-Factor
Definition	Information that only the user knows.	Something the user has.	Things that the user is.	Combination of two (or more) authentication factors.
Examples	Username, Password, Personal Identification Number (PIN)	Tokens, Public Key Infrastructure (PKI), Digital Certificates, Digital Signatures, X.509 Certificates	Biometrics	Password along with a token
Advantages	• Low cost. • Users are very familiar with them.	Enhance security without the need of hard-to-remember passwords.	• Ideal for accessing high-security systems. • Difficult to duplicate.	• Stronger security than if you used only one factor to authenticate. • Can be inexpensive.
Disadvantages	• Subject to attack. • Could be stolen. • Often forgotten. • Vulnerable to users that set easy-to-guess passwords. • Vulnerable to users that reuse their password on multiple accounts.	• Extra cost associated with a token. • Fraudulent use could allow improper disclosure of highly sensitive data.	• Need for complex additional hardware/ software to support biometric. • Potential high cost. • Potential to misuse the collected data. • Application obtaining data on the user that was not supposed to be accessible.	• Users can disable this feature. • Could create inconvenience for the user.

be tracked, and deviations of behavior may indicate a problem. The general idea of the approach can be seen in Figure 2 taken from (Jeong & Choi, 2012), which has proposed a method to provide security for mobile resources that reside in the cloud. This process begins with the user utilizing various systems/ sensors in order to enter his/her credentials. Then, this information gets processed through the means of several policies to determine the information/resources the users are allowed to access. *Risk-based authentication* consists of having different levels of authentication in a system and depends on the level of security the data that the user is trying to access has, and the way that the system policies are set up based on a risk score threshold. For example, suppose a system has been set up with a set of security policies. If a user is in a location that he/she frequently visits with his/her smartphone and is attempting to access high-level security data, then the system will utilize a basic method of authentication (such as a username/password combination) since the user was identified properly (the risk score was low). On the other hand, if the location of the user is not recognizable by the system and neither is his/her device, the system will opt to utilize more advance methods of authentication such as multi-factor authentication (the risk score was high).

Another biometric approach is *Keystroke Dynamics-based Authentication (KDA)* (Hwang, Cho, & Park, 2009; Changa, Tsaib, & Lina, 2012), which is a behavioral biometrics and can be seen as an approach used to analyze the way that a person types text on a keyboard/touchscreen. KDA can be utilized as part of a mobile user authentication method. For example, a user is trying to get authenticated in a mobile application and for this, he/she needs to enter his/her username/password combination. In this case, in addition to verifying the correct input, the system will analyze the way that the user entered the data based on factors such as the time he/she took to enter the credentials, and any other patterns found

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Platform for Authentication service.



using the smartphone's sensors. The work of Hwang et al. (2009) evaluated the way that KDA could be utilized in mobile devices via a framework from their paper, as shown in Figure 3. The framework contains three phases: the enrollment phase, the classifier phase, and the authentication phase. In the authentication phase, the user enters his/her credentials and the keystroke pattern gets analyzed. If the distance between the prototype and the pattern is lower than an established threshold then the user is legitimate, otherwise, an impostor. An interesting effort also in the KDA seeks to offer a more generalized typing authentication based approach that leverages biometric information that can be gleaned from the actual typing habit of the user (Feng, Zhao, Carbunar, & Shi, 2013), which would be stored electronically. The approach, called Typing Authentication and Protection (TAP), has two stages, which are shown in Figure 4: at logon, there would be an analysis to match the typing with the expected pattern of the user, while, after logon, the typing behavior of a user would be continuously checked with a variation from the expected having the potential to raise an alert.

The next approach, a *voice biometric* (Authentify, 2006), is realized as a numerical representation of the sound, pattern, and rhythm of an individual's voice, and is considered to be a secure approach due to the fact that, similar to fingerprints, such patterns are unique to each individual. To use a voice biometric, a user must perform a two-step process of enrollment and authentication, as shown in Figure 5, and taken from (Li, Tian, Yang, Shi, & Zhang, 2009). When a user enrolls, his/her biometrics are stored in the database, e.g., a voiceprint for the voice biometric. At a later point in time, whenever a user attempts to access a mobile application, in order to authenticate himself/herself, he/she needs to speak a phrase. The mobile device captures this phrase, extracts its features (such as sound and rhythm), and tries to match these with the voice prints/features that are stored in the database for that particular user. If the features stored in the database are similar to the features obtained from the user to a certain extent (higher than a threshold), then the user can access the application, if not, then he/she will be denied access. Another

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Three steps of the KDA framework.

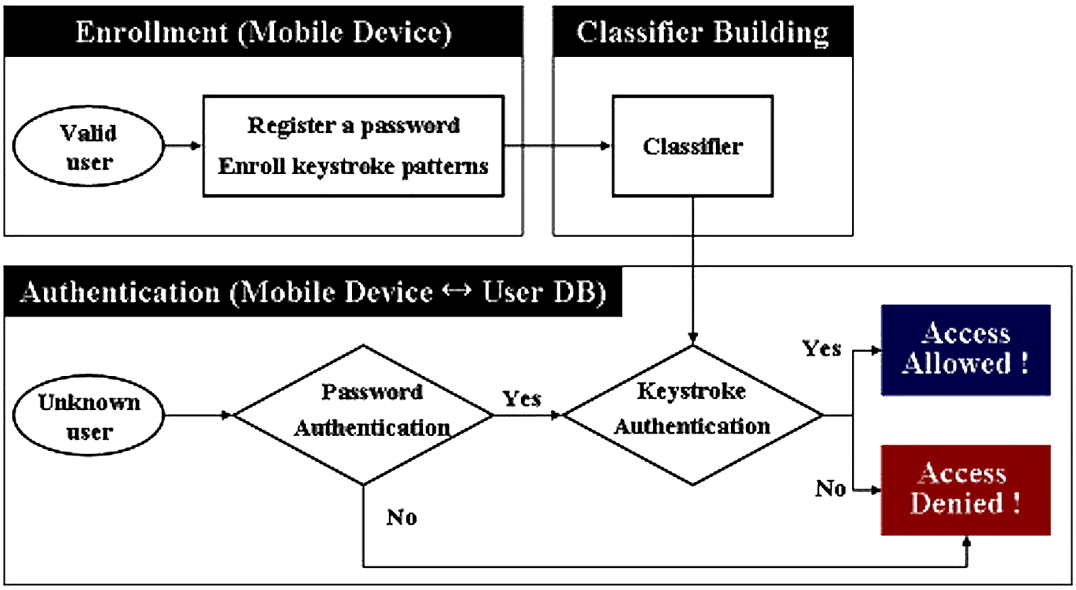
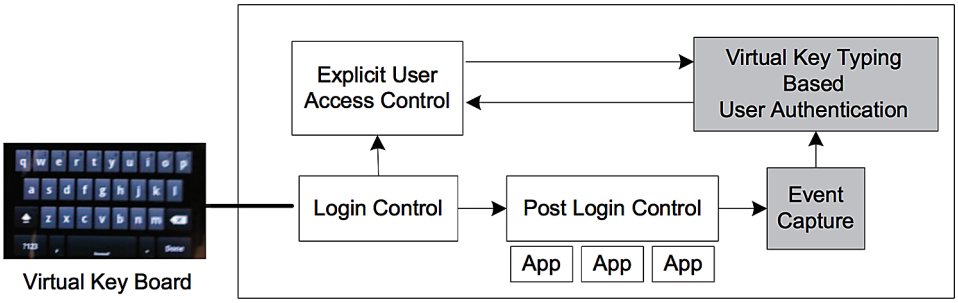


Figure 4. Two stages of the TAP framework.



research work related to the voice biometric approach in (Baloul, Cherrier, & Rosenberger, 2012) utilizes speaker recognition techniques to create a challenge-based method for mobile authentication in order to avoid security threats such as replay attacks.

Lastly, there has been an effort that considers the role of biometric when utilized within a *multi-factor authentication* context. Such a framework proposed in (Siddiqui, Abdullah, Khan, & Alghamdi, 2014) is comprised of a three-factor cloud-based user authentication method. The framework contains four phases: a registration authority phase, a smart application activation phase, a login phase, and a password phase. Authentication occurs in the login phase, where the user enters his/her username/password combination along with his/her fingerprints with the use of his/her mobile camera. Once these credentials get validated, if they are correct, then the system will send a one-time password (OTP) to the user's mobile device based on the calculation of a timestamp for further authentication. There are two ways that the system can reject the user at this point: if the timestamp sent with the OTP has an incorrect format, or if the result of the current timestamp minus the timestamp sent by the user is greater than an expected

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

interval. Otherwise, if the user entered the right OTP along with a valid timestamp, then the system will grant access to the user. This process is shown in Figure 6 as excerpted from Siddiqui et al. (2014). Another effort of note (Aronowitz et al., 2014) is very ambitious in combining a policy manager along with three different and sophisticated biometric techniques in one multi-model approach: *chirography*, which is based on an analysis of the touch of a user on a mobile device screen that is combined with both facial and voice recognition. The block diagram for the approach is shown in Figure 7.

AUTHENTICATION IN MOBILE HEALTHCARE APPLICATIONS

This section presents several mobile healthcare applications focusing on the respective methods of authentication to provide access to a user. The objective of this section is to evaluate the authentication

Figure 5. Enroll model and matching of biometric system

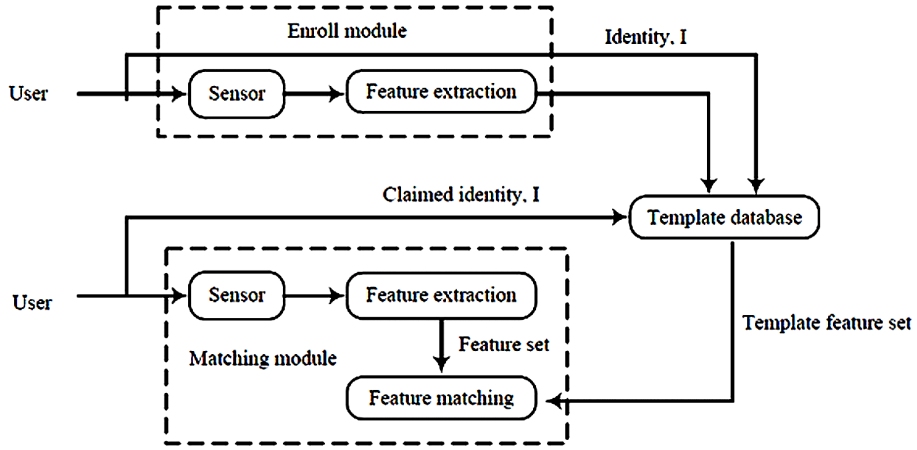
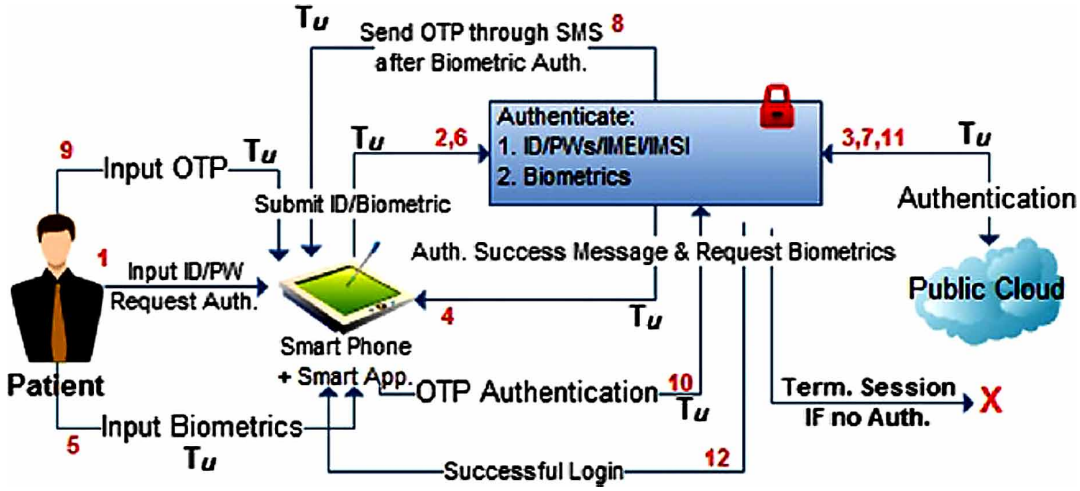
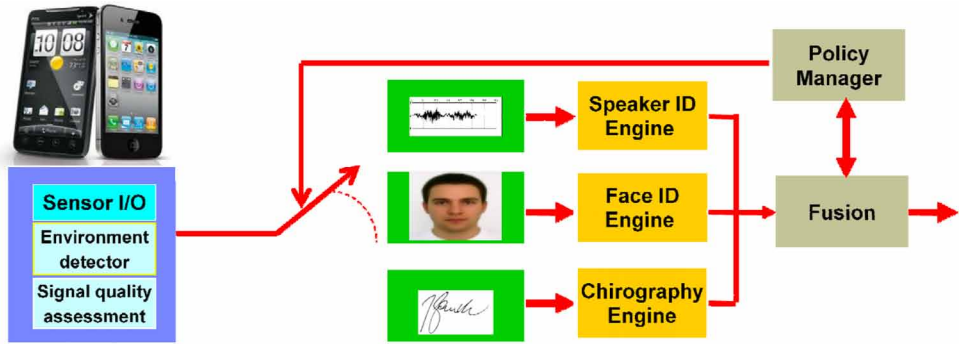


Figure 6. TMIS patient login authentication workflow



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Multi-modal biometrics authentication system



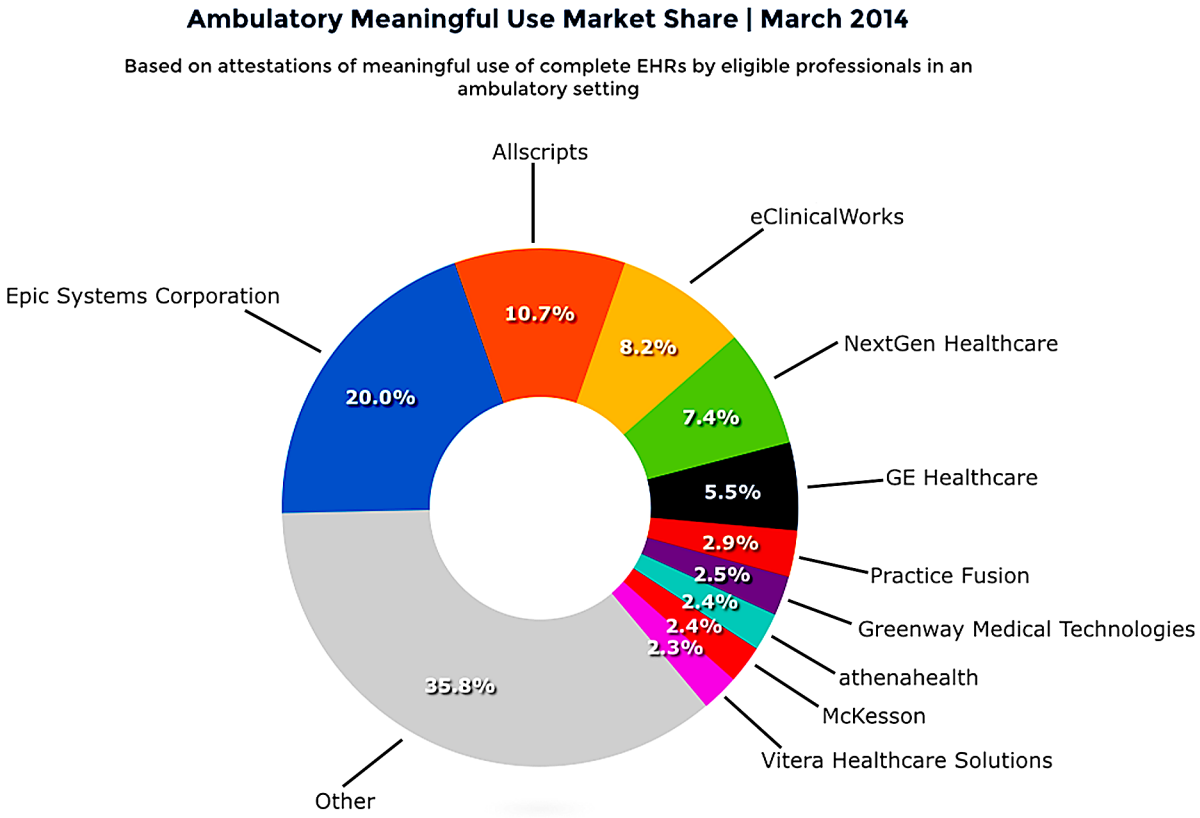
methods mobile healthcare applications utilize. The HIT market, in particular regarding electronic health records (EHRs), is diverse and extensive, with a reported 179 vendors supplying EHRs to 4,634 hospitals (Health IT, 2015) but is dominated by 10 products (McCormack, 2014), as shown in Figure 8. For the purposes of this paper, we focus on the "Top Ranked Virtualized & Native Mobile EHR Applications" (Black Book Market Research firm, 2014). From this list of top EHRs, we have selected the following representative set that was chosen based on their usage of advanced authentication techniques: drchrono, Allscripts Wand for Professional EHR, Kareo, Practice Fusion, and Greenway Healthcare PrimeSUITE.

The EHR drchrono (drchrono Security, 2015) is a collection of web-based, iPad, and Android tools with strong security capabilities that allow medical providers to perform tasks associated with viewing medications, allergies, and chart history of their patients and can be utilized to add patients, schedule appointments, among other features. In addition, patients that arrive at the medical practice can be given iPads for patient check-in, replacing the clipboards that are historically utilized. drchrono is HIPAA compliant (Online Tech, 2011) and meets industry-standard security practices, and authenticates users by utilizing a username/password combination along with a code that the user has to enter every time he/she logs in. This achieves *two-factor authentication* via the Authy app (Authy, 2015), where a combination of two factors is employed: a username/password (what a user knows) in conjunction with the mobile device itself (what a user has). Such a code is sent to the user's mobile phone and changes every 20 seconds. The app for drchrono also utilizes the auto-logoff feature, which automatically logs off users that are logged into the account but have been inactive for a certain period of time. In addition to this, drchrono utilizes a digital certificate in order to verify that the user is authenticated correctly and is in the correct site. Two-factor authentication is one of the key technologies that can provide an additional layer of protection to sensitive health care data (drchrono, 2013). Note that drchrono now offers fingerprint scanning on Apple mobile platforms to further control access to information. Lastly, there is the mobile application "drchrono EMR/ EHR" (drchrono Inc, 2015) for iOS platforms. The various drchrono products and applications are all certified stage 1 Meaningful Use criteria. Programmatically, drchrono provides a rich API and SDK to allow third party applications to be built that leverage the security capabilities while offering new ways to access health information.

The second EHR, Allscripts Wand for Professional EHR (Allscripts, 2014), provides a robust EHR with mobile anytime access, clinical decision support, and to track for patients during a visit, health reminders, medication management, potential allergies/interactions, ordering labs, procedures, medications, and so on. From a mobile perspective, Allscripts Wand offers a mobile client for both iPads and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Market Share of EHRs



iPhones, providing the medical provider with secure access to the Allscripts Professional EHR from remote locations. The corresponding Allscripts Remote app offers similar capabilities for Android devices. Like drchrono, Allscripts offers the ability for biometric two-factor compliant authentication through the BIO-key company (BIO-key, 2013) augmenting the username/password combination with fingerprint scanning and recognition. Allscripts Wand is intended to give medical providers flexibility through the usage of an iPad that travels with the physician as he/she moves around the medical facility to check on active patients, prep for new patients, and handle access when outside of the medical practice (nights, weekends, etc.).

The third EHR is Kareo (2015), which is available as a full-featured web-based and cloud-based solution whose 2014 edition is certified for Meaningful Use stages 1 and 2. Kareo provides both web-based, and iPad and iPhone based capabilities that support: patient registration/scheduling; a dashboard for medical providers with a schedule of the day's appointments, ability to approve prescription refills, review lab results, etc.; and a patient chart, documentation, and ordering functions for the actual appointment. At this time, patients are only provided with a web-based patient portal. Kareo was designed by doctors for doctors to be intuitive and provides features such as note-taking capabilities, one-click access to a patient's chart, shortcuts for the top 120 primary care conditions, while ensuring that all of the clinical information presented in the EHR is reliable and can be accessed fast. As for security features, Kareo has automatic backups of data and it provides customers with the option for them to setup their own password complexity policy for the patient portal. This meets PHI and HIPAA compliance and also allows

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

permissions to be set on a patient-by-patient basis to protect patient privacy. In addition, if a user tries to login many times and is unsuccessful, the Kareo patient portal will lock out the user's account until administrative action can be taken. At the present time, Kareo does not support two-factor authentication.

The fourth EHR, Practice Fusion, like Kareo, is a web-based cloud computing solution that is both web-based and iPad-based, but not currently supporting iPhone and Android. Practice Fusion has features such as: electronic prescribing, medical charting, scheduling, unlimited support, a personal health record for patients (Practice Fusion, 2009), and is HIPAA and Meaningful Use stage 1 and 2 compliant (Practice Fusion, 2013). According to Practice Fusion, the EHR contains solid user authentication and utilizes user login time windows to determine when a user logged on to the system, automatically logs out the user if he/she has not been utilizing the application for a certain amount of time, and enforces the use of strong login passwords. Also, Practice Fusion contains well-defined user roles and access levels. In addition to this, the data that belongs to each of the practices is kept separate for each organization in order to maintain privacy and confidentiality. The data sent between the users and the Practice Fusion servers is encrypted during transmission using SSL/HTTPS, and utilizes ultra-secure tokens (64-byte hexadecimal strings). For the newer iPads, fingerprint authentication for enabling the mobile device is supported, but the technology has not been integrated as a two-factor authentication with Practice Fusion.

The final EHR, Greenway Health PrimeSUITE (Greenway Health, 2015), is also a web-based and cloud-based solution that is HIPAA and Meaningful Use stage 1 and 2 compliant. In terms of mobile support, PrimeMOBILE offers the ability for medical providers to utilize Apple, Android, or Windows devices, making it among the most platform-versatile EHRs. However, the patient portal that is provided, PrimePATIENT, is web-based. Note that while Greenway PrimeSUITE does not yet support two-factor authentication, third party hosting vendors can provide that capability (iTELAGEN Healthcare IT Redefined, 2015). In fact, one vendor offers PrimeSUITE access through any computer or mobile platform/device that provides tunneled SSL as well as two-factor authentication (MDS Medical, 2014).

The mobile healthcare applications we chose as part of our analysis and, that are within the top ten EHRs, provide different security mechanisms to authenticate users. Table 2 depicts some of the main factors we are evaluating to consider whether an application is highly secure or not. These factors are: the authentication mechanism used, the policies/methods utilized to keep the credentials secure and, if it contains an auto log off feature, which we consider its crucial in environments where many users access a mobile application on the same mobile device. In the Background and Motivation section, we discussed the three main authentication factors and the way that two or more of these could be combined in order to have a two-factor/multi-factor authentication approach. Even though this type of mechanism is becoming increasingly popular in systems that contain highly sensitive data, only two of the EHRs evaluated (drchrono and Allscripts Wand for Professional EHR) contain this feature (note that Greenway Healthcare PrimeSUITE relies on third-party vendors to have two-factor authentication). The authentication method commonly utilized in the evaluated mobile applications is the username/password combination mechanism. While some of the EHRs utilize this mechanism in combination with another authentication mechanism to provide two-factor authentication (drchrono and Allscripts Wand for Professional EHR); others do not provide a second authentication factor and provide password policies (Kareo) or enforce the use of strong passwords (Practice Fusion) instead. We can also notice that biometric methods are utilized in three of the EHRs we presented (drchrono, Allscripts Professional, PrimeSUITE). What all the EHRs had in common in terms of securing the authentication process was the fact that all of them include the auto log off feature. This is strongly required in mobile applications that contain highly sensitive data since it logs out the user after a certain period of inactivity in the session.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Main Authentication features in evaluated mobile applications

	Username/ Password	Password Enforcement/ Policy	Digital Certificates	Biometric Authentication	Two-factor Authentication	Automatic log off
drchrono	✓	✓	✓	✓	✓	✓
Allscripts Wand for Professional EHR	✓			✓	✓	✓
Kareo	✓	✓				✓
Practice Fusion	✓	✓				✓
Greenway Healthcare PrimeSUITE	✓			✓	✓	✓

While all the evaluated EHRs contain authentication mechanisms to provide a secure mobile application for users, each one of them has drawbacks. For instance, all of these have an option that involves a username/password combination. Some have mechanisms to ensure that the password is strong (drchrono, Kareo, and Practice Fusion), while others suggest users to have a strong password but do not enforce this requirement (Allscripts Professional and PrimeSUITE). In addition, the EHRs that enforce passwords do not offer two-factor authentication. Thus, if the user's password is guessed, then there is no other method to ensure that the user who entered the password is who he/she says he/she is.

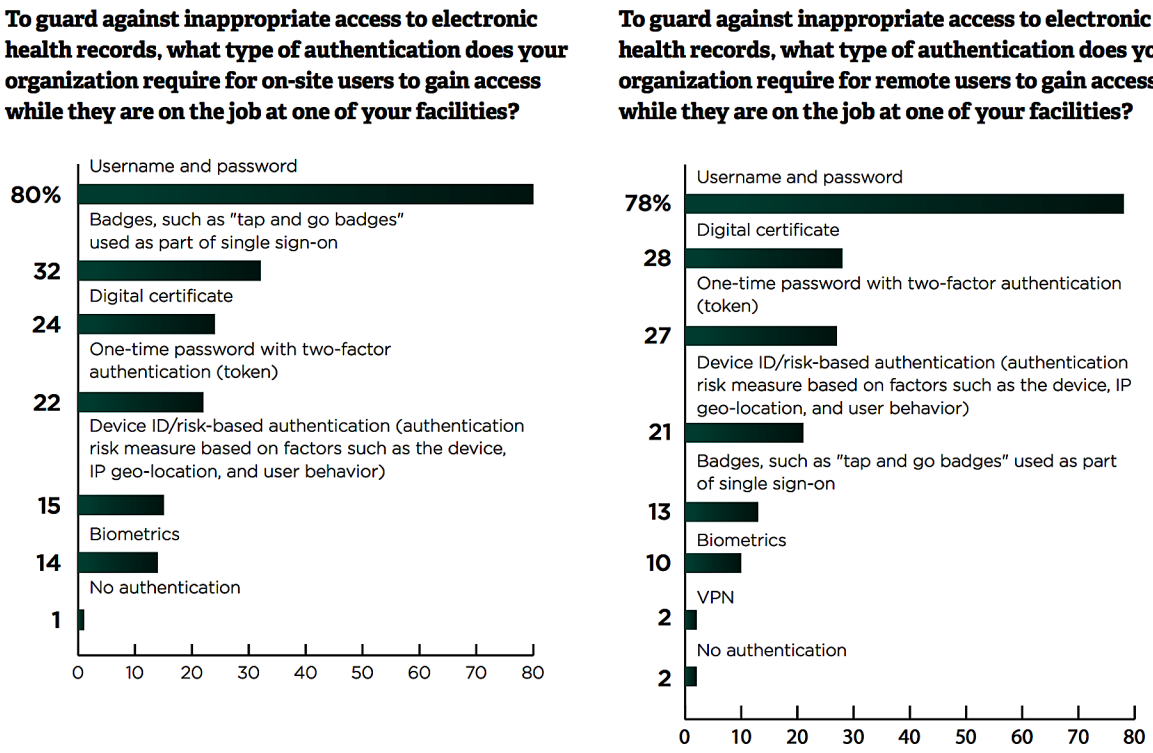
To complete the discussion, we highlight the *Healthcare Information Today* survey (Information Security Media Group, 2015), which contains answers to various questions on the way that healthcare entities manage regulatory compliance to the steps that are been done in order to secure health data. One question of particular note is relevant to the main purpose of this chapter, related to the authentication mechanisms healthcare organizations utilize to protect EHRs. The left side of Figure 9 shows (in percentages) what are the authentication mechanisms users utilize when they are inside the healthcare facility they work in and, the right side of Figure 9 shows (in percentages) what are the authentication mechanisms users utilize when they are outside the healthcare facility they work in and need to access EHRs. By looking at the results shown on Figure 9, we can see that the dominant method to secure EHR data is a username/password combination while methods that can be more suitable to protect this data (e.g. two-factor authentication and biometrics) are not generally used.

AUTHENTICATION REQUIREMENTS IN MOBILE COMPUTING

The objective of this chapter is to enumerate and explore the requirements for user authentication as it relates to mobile computing. We have identified some requirements on the critical areas that have to be addressed in order to achieve secure user authentication in support of the diverse needs of mobile computing, particularly as related to high security concern areas such as health care. In order to determine these, we took into consideration the limitations of the evaluated mobile applications in the previous section as well as the concerns expressed in the work presented by other researchers and that we briefly discussed throughout the chapter. There are five specific requirements identified to date, including: long-term mobile application sessions; beyond two-factor authentication; advanced threats and hacking;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Dominant user authentication methods in EHRs



multiple and simultaneous authentication; and, lost or stolen device. Each is reviewed in turn. To complete the discussion, these five requirements are juxtaposed into a framework that demonstrates the way that the requirements interact with one another via a three-phase process of device security, authentication process, and application security.

The first requirement, *long-term mobile application sessions*, deals with the unique nature of mobile applications in terms of the length and duration of a session under the control of a username/password logon. When one starts a session with a mobile app, and provides his/her username/password credentials, most apps assume that the user wants to store that information within the app so that subsequent logons no longer require providing that information. Some applications that involve payment or other charges require the user to re-authenticate often with just a password re-entry in order for the purchase to proceed or even just an acknowledgement to use a particular account, which utilizes the stored prior username/password combination. In fact, many applications retain the username/password information for long periods of time, and even remember this information after the application is closed and reopened. From a mobile authentication perspective, this could be a potential problem, particularly if a device is lost or stolen. Suppose that a user has a pharmacy app, a personal health record app (e.g., Microsoft HealthVault or Web MD), and a patient portal app (e.g., the Kareo patient portal from the prior section), all of which have HIPAA sensitive information on current prescriptions (pharmacy app), personal health data (PHR), and pending appointments or ordered medical tests (portal). If these apps do not re-authenticate, then they would be vulnerable to both access to PHI as well as potential malicious misuse if the mobile device is lost or stolen. For health care and other sensitive mobile applications, there is a need for the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

long-term mobile application sessions to be curtailed in some manner, perhaps timing out the application akin to the way that sessions are terminated in browsers, or by frequent or period re-authentications. The recommendation that we make for this requirement is to suggest that health care and other sensitive mobile applications remove the long-term storage of username/password in an app, so that each time the application is closed, this information is erased.

The second requirement, *beyond two-factor authentication*, involves the need to perhaps go beyond the concept of two-factor authentication in order to arrive at an even stronger process to authenticate, particularly in regards to health care and other sensitive mobile applications. The area of multi-factor authentication is targeting the need to go beyond just two factors. One of the most interesting aspects of mobile computing is the fact that the devices continue to have more powerful processing power, high resolution, advanced cameras and features, and additional capabilities via a wide range of sensors. There is an opportunity to exploit all of these features with authentication techniques that are compute-intensive such as encryption or utilizing the camera and processing for retinal scanning/recognition or facial recognition as a means to improve the security. The recommendation that we make for this requirement is that there should be a movement towards multi-factor authentication with the intent to take advantage of the growing computing and processing capabilities of mobile devices for more secure authentication. An example of an approach that goes beyond two-factor authentication is demonstrated in Siddiqui et al. (2014) as shown in Figure 6.

The third requirement, *advanced threats and hacking*, has the need to be aware of the widespread attacks that are underway against all types of computing infrastructure across a wide range of domains. In the health care area, one of the major health insurers in the United States, the health insurer Anthem, was the target of a major cyber attack in February 2015 (Anthem, 2015) and there have also been attacks on major retailers (Target and Neiman Marcus) and financial institutions such as JP Morgan Chase (Finkle & Hosenball, 2014). In addition, in the health care domain, there was a recent attack at the Children's Hospital in Boston, MA, and such an attack has the potential to directly impact patient care (Farrell & Wan, 2014). Further, there is a major assessment at the national level that has been undertaken (Clapper, 2013). There has also been a marked increase in mobile cyber attacks (Ruggiero & Foote, 2011; RT, 2014). In terms of mobile computing, these advanced threats and hacking have the potential to disrupt the way that we utilize mobile devices in all of our activities, with a clear impact on health, e-tailing, and financial applications. These attacks are nefarious since they don't involve a direct assault on the information on a mobile device, but instead target sensitive PHI and PII that are stored in the systems mobile applications are interacting with. As a result, our recommendation for this requirement is that there needs to be a realization that the systems that mobile applications utilize be more secure for these indirect attacks. For instance, the mobile application could enforce a multi-factor authentication approach as the one mentioned in the previous requirement (Siddiqui, Abdullah, Khan, & Alghamdi, 2014)

The fourth requirement, *multiple and simultaneous authentication*, focuses on the need for user authentication, in which a user of a mobile device is attempting to access a set of related systems through a single application. For example, consider a mobile application for a patient portal for one of the medical systems in the prior section of this chapter, that allows a user to do a number of simultaneous tasks including: request a prescription refill or its status which interacts with the HIT systems of a specific pharmacy; schedule an appointment for a blood test which interacts with the HIT system at a particular testing laboratory; review past medical examinations and data such as blood pressure, cholesterol, etc., that accesses your data in the EHR of your medical provider; and, store and access fitness data stored in the HIT systems of a fitness vendor. A mobile application with such breadth requires username/password

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

authentication with multiple systems either simultaneously or in some sequence. In such a situation, does the authentication check get sent to multiple systems in parallel? Or in a sequence where it has to be successful at one system before authenticating the next system in a chain? From a mobile computing perspective, the issue is the appropriate way to handle the need for multiple HITs authentications in a simultaneous manner with the same credentials. One danger in automating such an authentication process would be that if the mobile device has been taken over (by either another user or via an attack), such a chained authentication process would mean that multiple systems accessible to the mobile application would all be vulnerable. The recommendation for this requirement is to in fact mandate that multiple and/or simultaneous authentications be severely limited, even at the expense of aggravating the end user, in order to insure that the potential exposure of PHI and PII are limited.

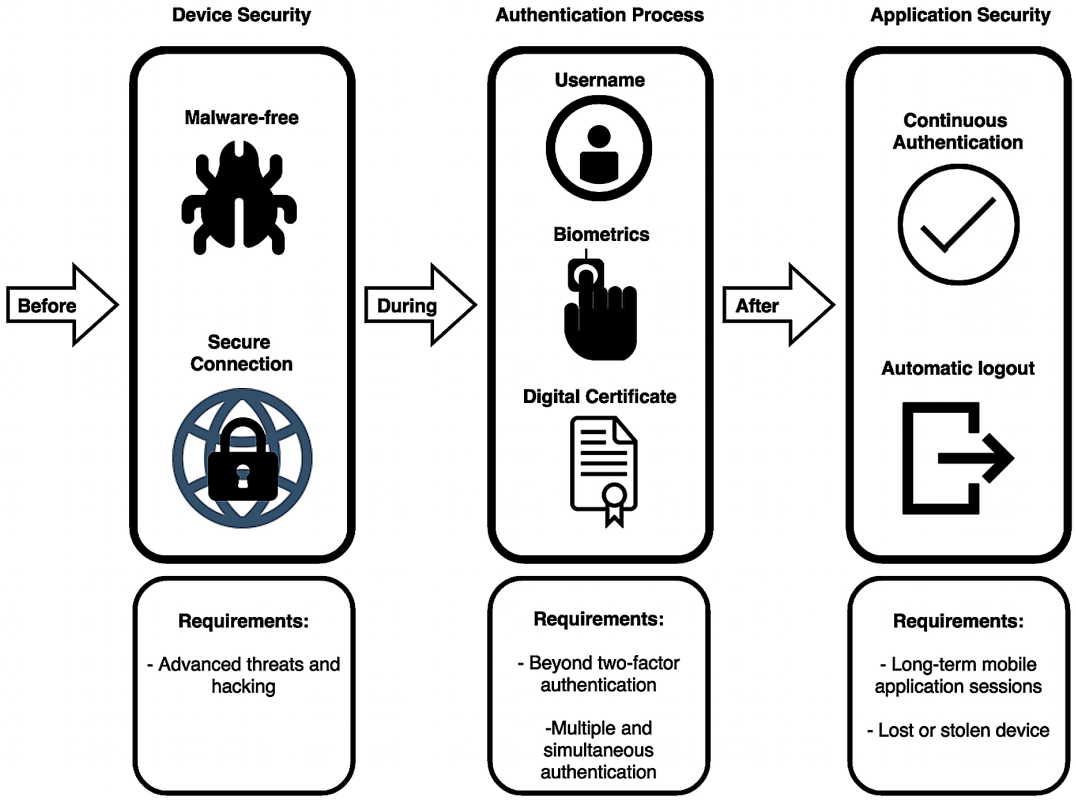
The final requirement, *lost or stolen device*, considers the case where a user's device is misplaced or stolen. This could affect many applications whose sessions are running at the time if the device does not have a strong authentication method (e.g. has an easy-to-guess pattern to unlock the device) or doesn't have an authentication method at all. For example, consider the case in which a person is utilizing his/her device to view his/her medical record and that even when the application is closed by the user, it remains in session for 10 minutes before the user is automatically logged out. If the device lacks a proper lock and the application is still running, then the person who stole the device could gain access to the user's confidential information, and maybe, obtain the permissions needed to modify or corrupt the stored PHI. This clearly indicates a strong requirement to insure that the user's private data will not be compromised in this situation. Taking this into consideration, our recommendation would be to implement a more secure mobile authentication approach to unlock a device. A research work that attempts to secure a mobile device (Sun, Zhang, Zhang, & Zhang, 2014) proposes a two-factor authentication approach, which consists of a user drawing on arbitrary regions on the touchscreen of his/her smartphone in order to unlock it, with the user being authenticated based on the geometric properties of the curves he/she drew, in addition to his/her behavioral and physiological characteristics. Our recommendation for this requirement is to consider that approach as well as to leverage other multi-factor and simultaneous approaches that that compromising PHI or PII can be mitigated when devices are lost or stolen.

The five requirements as presented in this section can be collected into a framework that can be applied to mobile applications that contain highly sensitive data to provide a secure authentication process. This is illustrated by the framework in Figure 10 which has a three phase process of device security, authentication process, and application security, shown from left to right; note that under each phase, the relevant five requirements per phase are indicated. The *device security phase* consists of the requirements utilized before user authentication takes place. In this phase, when the user opens the application and before he/she is given permission to log in, the mobile application will verify if the mobile device does not contain malware and if it's connected to a secure network (e.g., WAP or WEP connections). The device security phase satisfies the third requirement previously mentioned (*advanced threats and hacking*).

The *authentication process phase* contains the requirements that are used during the user authentication process. This phase covers the authentication process of the mobile application and consists on a secure authentication approach such as multi-factor authentication (depicted in Figure 10), which consists of the user's username, of a user's biometric trait (e.g., fingerprint, voice recognition) and, of a digital certificate that contains information about the user's device and determines if the mobile application is trusted. In addition, the authentication process phase supports the case when a mobile application connects to more than one system to retrieve information that is displayed in it, utilizing the concept of single sign-on (SSO), which consists of granting access to several systems with only one set of credentials. The

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Proposed process involving requirements



authentication process phase covers the second and the fourth requirements described (*beyond two-factor authentication* and *multiple and simultaneous authentication*). After the authentication process phase completes we need to keep verifying if the user is who he/she says he/she is. This is accomplished in the *application security phase* for the requirements that are applied when the user is authenticated. This phase handles this situation by continuously verifying the user's identity with methods such as KDA. In addition to this, the application security phase contains the automatic log off method, which consists of terminating the session of the mobile application if the user is inactive for a certain period of time or if any anomalies are detected in the user's behavior while he/she is utilizing the mobile application (e.g., change of keystroke patterns, sudden change of locations, sudden change of actions, etc.). Once the application logs out, the personal information of the user will be erased from the mobile device (if any was stored during the session). This final application security phase satisfies the first and fifth requirements described (*long-term mobile application sessions* and *lost or stolen device*). Note that the methods presented in each of the phases could be modified in order to provide different security methods.

FUTURE TRENDS

This section identifies and reviews approaches that are emerging and may impact authentication in two different areas. First, we explore the emerging technologies for biometric and multi-factor authentica-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tion, as applied to mobile platforms, which includes continuous authentication in order to keep verifying the user's credentials after he/she gets authenticated and while the session is active. Then, we review the new mobile operating systems for mobile devices that include Firefox OS, Tizen, Ubuntu Touch for Mobile, and Jolla Sailfish. Over the past few years, a general consensus has coalesced towards the likely predominance in the very near future of widespread adoption and usage of biometrics authentication and multi-factor authentication for user authentication in mobile devices/applications (Whittaker, 2013; Beranek, 2014; Reisinger, 2014; Coplan, 2014; Zaibak, 2014). In addition to this, in the Background section of this chapter, we reviewed several research efforts (Siddiqui, Abdullah, Khan, & Alghamdi, 2014; Jeong & Choi, 2012; Maiorana, Campisi, González-Carballo, & Neri, 2011) that evaluate biometric and multi-factor authentication approaches and propose ways in which they could be implemented to obtain a more secure and user-friendly mobile user authentication scheme than what the market offers these days. Our objective in this section is to explore these methods in greater detail and their application to mobile computing in the future from a more technological perspective.

Biometrics methods have begun to be adopted for mobile devices. Mobile network maker Ericsson conducted a survey, which consisted on asking 100,000 people of over 40 countries what were their thoughts about biometric smartphones. Seventy-four percent of respondents said that such device would "become mainstream" throughout 2014 (Beranek, 2014). We can say this prediction became fairly true in September 2013 when Apple released the iPhone 5s, which includes a fingerprint reader with the purpose of modifying the way users authenticate to the system and/or to applications, and to improve user usability. In addition to this, in April 2014, Samsung released the Samsung Galaxy S5, which contains a fingerprint scanner as part of the security package of the mobile device. Further, on May 2014 Samsung Senior Vice President Rhee In-jong reported that his company is exploring ways to increase the utilization of biometrics in their mobile devices to incorporate "iris detection" into the smartphones that will scan a user's eye to determine if the iris matches what the device has on-file (Reisinger, 2014). Also, besides fingerprint scanning and "iris detection", other biometric factors that are currently being considered are facial recognition, voice biometric, and vein. One effort (Beranek, 2014) postulated that the biometric method that will likely be utilized on upcoming authentication schemes would be voice biometrics since it is considered to be less invasive than iris detection and fingerprint scanning. In addition, in 2014, CA technologies joined the Fast Identity Online (FIDO) Alliance in order to create a novel authentication model for mobile applications (Coplan, 2014). Instead of seeing authentication as a one-time event, this effort seeks to implement a continuous authorization system based on dynamic contextual-based intelligence. The approach is that a system needs a certain level of intelligence in order to generate a risk profile, and needs to be able to coordinate the authentication process over a session. The approach would eliminate password-based mobile authentication to be replaced by techniques that employ context and behavioral analytics, and combine multiple sources of credentials and intelligence that triangulates context against risk. Currently, the literature suggests that continuous authentication will become one of the preferred methods to do user authentication in mobile applications that contain highly sensitive data (Qualcomm, 2015; Zeltser, 2015). The main idea behind such a method is to continuously verify the user's identity in a transparent way (without interrupting the user's session) by looking at the interactions that the user has with the application. For this, we can utilize biometrics, keystroke patterns, and sensors that the smartphone contains (e.g. accelerometer). If the mobile application detects an anomaly in the context that is being evaluated, then it could ask the user to validate his/her identity or, it could close the session.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Concurrent with the aforementioned biometric authentication, there are emerging technological solutions that have begun to appear for use with mobile devices. First, the two major mobile platforms, Android and iOS, are offering multi-factor authentication (MFA) and biometrics authentication applications in order to serve as an additional method of authentication for users to gain access to some of their accounts, namely, "Google Authenticator (GA)" (Google, Inc., 2015), and "Eye Verify Biometric Data Tool" (EyeVerify LLC, 2015). All of these mobile applications are available on both Android and iOS platforms: the first one serves as a multi-factor authenticator, and the second one serves as a biometric authenticator. Basically, these mobile applications add another level of security to user accounts that support such authentication methods. In addition, there are more global approaches to mobile security that apply to multiple devices. For example, Symantec, a world leader in PC-based, cloud, and mobile security solutions offers Validation & ID Protection (VIP Access (Symantec, 2013)) for all mobile platforms (iOS, Android, Windows, Blackberry, etc.) that provides a unique security code for when logging onto participating online financial and merchandise web sites. VIP works for both mobile and desktop devices and is available on the stores for the aforementioned platforms. Similar to Symantec's approach, VASCO (2015) a leader of providing two-factor authentication and electronic signature solutions to financial institutions, has created DIGIPASS for Apps (VASCO, 2014), which is available on almost any mobile device platform (includes major platforms such as iOS, Android, Blackberry, and Windows). This is an SDK that provides enhanced security for mobile applications by allowing users to establish two-factor authentication (supports PIN and fingerprint biometrics features) and electronic signing into their mobile applications. Moreover, the SDK provides a risk-scoring feature, which consists of embedding platform and context elements into the authentication process in order to strengthen the security of the application.

Finally, we discuss several new operating systems for mobile devices. These are Firefox OS (Mozilla, 2015), Tizen (2014), Ubuntu Touch (Ubuntu, 2015), and Jolla Sailfish (Sailfish, 2015). All of these are open-source, Linux-based mobile operating systems, and they have many similarities as well as features that make them unique. Since all of these are relatively new to the mobile market, some lack detailed information on the way that user authentication is handled and have limited material on other relevant security features. Nevertheless, we provide a general idea of each system and its available security features.

Firefox OS was released in 2013 and is a Linux kernel-based open-source operating system available for smartphones, tablets and smart TVs. In terms of implementation, all of the applications are web applications; nevertheless, the application's pedigree or provenance determines in which category the web application is. The applications can be classified into three categories: *certified applications* that come preinstalled in the device; *privileged applications* that the application market has vetted meaning that they have passed through a process that checks that they cannot perform malicious actions; and *web applications*, which are basically webpages that are accessible through a mobile browser. The majority of third-party applications are in this last category and contain the least available permissions due to the fact that the marketplace does not revise the applications that fall in it. As for security, Firefox OS implements it on several layers (Gadyatskaya, Massacci, & Zhauniarovich, 2014). One of the layers provides application sandboxing using Discretionary Access Control (DAC), and another layer implements the permission system utilizing Mandatory Access Control (MAC). As a result, the application sandboxing prevents applications from having access to the data of other applications. To achieve this, each application contains a unique low-privilege user ID (UID). This approach can prevent attacks from malicious websites, but the amount of overhead that is created can be significant since for two applications that utilize the same data or share a great part of the data, each needs its own copy. As for permissions, the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

mobile OS can deny, allow, or prompt default actions. Similar to Android, an application must specify the permissions needed in a manifest file. In addition to these security features, Firefox OS enforces a content security policy, which is utilized to ensure that the content shown on the page is the one that is supposed to be there. This prevents attacks such as cross-site scripting. Moreover, the Firefox OS includes the capability of allowing an application to request two sandboxes, one for itself and one for any website it accesses, in order to prevent the application from attacks by a website's contents.

Tizen OS was released on 2012 as an ordinary Linux distribution. In terms of applications, it supports mobile website, hosted, and packaged applications. *Mobile website applications* open in a mobile device's web browser, *hosted applications* are client applications that store and access their resources on a remote server, and *packaged applications* are the ones whose have all of their resources in their package. The latter two applications can be published in the Tizen store, and packaged applications are the only ones that have access to the Device API. Similar to Firefox OS, each application contains a manifest file. However, instead of each application having a unique UID, there are only two UIDs: root and app. Therefore, a DAC mechanism cannot be utilized to separate app data (do sandboxing). As an alternative, Tizen OS utilizes the Simplified Mandatory Access Control Kernel module, which manages access to resources via simple rules in the form {subject, object, permission}. In addition to providing sandboxing, Tizen OS provides the ability for applications to interact with one another. Developers can establish private communication between their applications by utilizing application or data control mechanisms. Tizen OS refers to permissions as privileges, of which there are three types: *public privileges* which are available to all applications; *partner privileges* which are only for companies registered as trusted partners in the Tizen store; and, *platform privileges* which are only available for developers working for the Tizen consortium. The privilege level is identified by verifying the signatures that come with the application, of which there are two, a developer signature and/or distributor signature. The developer's signature can be utilized to grant additional permissions to applications that contain the same signature, and the distributor's signature is utilized to sign a system application in order to assure that the application is safe to run, and belongs to the Tizen store by default. Similar to Firefox OS, Tizen has cross-site scripting options for apps that want to restrict the sources of web content, and it can also limit the webapp navigation to a domain list by specifying such in the manifest file. Lastly, a unique security feature of Tizen OS is that the developer has the ability to encrypt the device's application resources (i.e. HTML, JavaScript) in order to protect such from plagiarism.

The next OS, Ubuntu Touch was released in October 2013, with an announcement in September 2014 that was recently released in Europe. In terms of data security for mobile applications, the Ubuntu Touch has an approach called Application Confinement (Ubuntu Developer, 2014) that makes sure that the applications only have access to the files that belong to them. In order to share content between applications, Ubuntu Touch provides a Trusted Helper service called the Content Hub that is still in development. In addition to this security feature, the OS has a trust model (Ubuntu Developer, 2013) to determine which applications can be trusted and which can't for purposes of granting permissions. The applications that are not trusted by the OS are the ones that run in a restricted sandbox, and applications that are published in their app store. These cannot access features such as user data, privileged APIs, sensitive APIs, or another application's data without the user's permission. The applications that are trusted by the OS are the ones that come installed with the device. Since these trusted applications are not under confinement, they have permissions such as: being able to access resources or data available in the user's session, and limited access to system services and data as defined by the OS. Permission to sensitive data can be granted or denied to applications published in the app store when a user is trying

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to install them. As a result, the user can be aware of which permissions the application needs to have access to and based on this, make a decision on whether to install it or not.

The last OS is Jolla Sailfish OS. Even though its release was made on November 2013, the operating system main page contains limited information related to its security features. Nevertheless, on March 2015, Jolla announced that it will be teaming up with SSH Communications in order to create a high-security version of Sailfish OS (Lomas, 2015). The version is targeted to be utilized by government officials, and for corporations that need a secure mobile solution. However, anyone who wants a more secure mobile solution for his or her personal use will also be able to acquire the product.

Finally, we note that there are several secure mobile solutions that are currently out in the marketplace. Two of these are the Blackphone (Silent Circle, 2015), and the Tor mobile operating system (O'Neill, 2014). Blackphone was launched on June 2014, and its operating system, PrivatOS, can be defined as a combination of a fortified Android operating system along with a set of applications designed to provide users absolute privacy. Tor was announced on February 2014 based on Tor's The Amnesic Incognito Live System (TAILS) approach, which is an anonymous operating system for PCs. The applications, which can be downloaded free with Tor, will provide strong anonymity to the user, and features such as encrypted voice, video, text messaging, and security features that can disable location tracking. The system will be available on stock Android and Ubuntu products.

CONCLUSION

This chapter has provided a detailed discussion of alternative user authentication methods with an aim towards identification and explanation of user authentication requirements for mobile computing, supplemented by a discussion of its potential usage in a healthcare application. In support of this work, the *Background* section reviewed the health care domain, its systems, and the usage of information, coupled with a discussion of a future potential usage of mobile devices and their applications, in general, and as related to health care. In terms of authentication, the *Security Concepts Impacting Mobile Authentication* section explained concepts such as user/password identification, PKI/certificates, biometrics, and multi-factor authentication. Using these two sections as a basis, the *Authentication in Mobile Healthcare Applications* section explored the specific authentication methods for actual health information technology systems (e.g., drchrono, Allscripts Wand for Professional EHR, Kareo, Practice Fusion, and Greenway Healthcare PrimeSUITE) that have requirements to simultaneously access and protect data. This set the stage for the discussion in the *Authentication Requirements in Mobile Computing* section, which proposed and explored five requirements (long-term mobile application sessions; beyond two-factor authentication; advanced threats and hacking; multiple and simultaneous authentication; and, lost or stolen device) that identify needed capabilities that currently transcend traditional authentication methods. These five requirements were then collected into a framework that demonstrates their interactions and had a three-phase process for: device security, authentication process, and application security. To complete the chapter, the *Future Trends* section identified approaches that are emerging and may impact authentication including continuous authentication, the increase of multi-factor authentication in mobile devices (Google and Apple), new mobile operating systems (Firefox OS, Tizen, Ubuntu, Sailfish), approaches for mobile security (like VIP Access), and secure mobile OSs (Blackphone and Tor).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Abe, T., Itoh, H., & Takahashi, K. (2007). Implementing identity provider on mobile phone. In *Proceedings of the 2007 ACM workshop on Digital identity management* (pp. 46-52). ACM. 10.1145/1314403.1314412

Allscripts. (2014). *Allscripts Professional EHR*. Retrieved from http://www.allscripts.com/products-services/products/professional-ehr/features

Anthem. (2015, May). *How to Access & Sign Up For Identity Theft Repair & Credit Monitoring Services*. Retrieved from https://www.anthemfacts.com/

Apple iOS 9. (2014). *Health, an innovative new way to use your health and fitness information*. Retrieved from http://www.apple.com/ios/health/

Aronowitz, H., Li, M., Toledo-Ronen, O., Harary, S., Geva, A., Ben-David, S., … Nahamoo, D. (2014). Multi-Modal Biometrics for Mobile Authentication. In *Proceedings of Biometrics (IJCB), 2014 IEEE International Joint Conference* (pp. 1-8). IEEE. doi: 10.1109/BTAS.2014.6996269

Authentify. (2006, December 28). *Voice Biometric Authentication*. Retrieved from www.authentify.com/solutions/voice_biometrics.html

Authy. (2015). *Authy*. Retrieved from https://www.authy.com/

Baloul, M., Cherrier, E., & Rosenberger, C. (2012). Challenge-based speaker recognition for mobile authentication. In *Proceedings of International Conference of the Biometrics Special Interest Group (BIOSIG)* (pp. 1–7). Academic Press.

Beranek, B. (2014, February). *Biometrics on the smartphone: The future of mobile authentication*. Retrieved from http://whatsnext.nuance.com/customer-experience/biometrics-smartphone-future-mobile-authentication/

Bhargav-Spantzel, A., Squicciarini, A. C., Bertino, E., Modi, S., Young, M., & Elliot, S. J. (2007). Privacy preserving multi-factor authentication with biometrics. *Journal of Computer Security*, *15*(5), 529–560.

BIO-key. (2013). *Biometric Fingerprint Technology for Allscripts*. Retrieved from http://www.bio-key.com/products/data-sheets/allscripts

Black Book Market Research Firm. (2014, June). *Black Book Rankings 2014 Survey Results – Top Ranked Virtualized & Native Mobile EHR Applications*. Retrieved from http://blackbookmarketresearch.com/wp-content/uploads/2014/08/iPad-EMR_EHR-Apps-JUNE-2014-BLACK-BOOK-HCINFO.pdf

Boxall, A. (2013, January). *War for Your Pocket: These 5 New Operating Systems Plan to Battle Android and iOS in 2013*. Retrieved from http://www.digitaltrends.com/mobile/five-new-mobile-operating-systems-ready-for-launch-in-2013/

Centers for Medicare & Medicaid Services. (2015, October 16). *Medicare and Medicaid Programs; Electronic Health Record Incentive Program-Stage 3 and Modifications to Meaningful Use in 2015 Through 2017*. Retrieved from https://www.federalregister.gov/articles/2015/10/16/2015-25595/medicare-and-medicaid-programs-electronic-health-record-incentive-program-stage-3-and-modifications

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Changa, T., Tsaib, C., & Lina, J. (2012, May). A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices. *Journal of Systems and Software*, *85*(5), 1157–1165. doi:10.1016/j.jss.2011.12.044

Clapper, J. R. (2013, March). *Worldwide Threat Assessment of the US Intelligence Community, US Senate*. Retrieved from http://www.intelligence.senate.gov/130312/clapper.pdf

CMS.gov. (2015). *Electronic Health Records (EHR) Incentive Programs*. Retrieved from https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html?redirect=/ehrincentiveprograms/

Codjia, M. (2010, October 15). *Define Medical Practice Management System*. Retrieved from http://www.ehow.com/facts_7151485_define-medical-practice-management-system.html

Coplan, S. (2014, May). *CA joins the FIDO Alliance to change the future of user authentication*. Retrieved from http://blogs.ca.com/2014/05/15/ca-joins-the-fido-alliance-to-change-the-future-of-user-authentication/

drchrono. (2013, August 13). *drchrono Offers Unmatched Protection Against Electronic Medical Record Security Threats*. Retrieved from http://www.prnewswire.com/news-releases/drchrono-offers-unmatched-protection-against-electronic-medical-record-security-threats-165966406.html

drchrono Inc. (2015). *drchrono EHR/EMR*. Retrieved from https://itunes.apple.com/us/app/drchrono-ehr-emr/id369191782?mt=8

drchrono Security. (2015, February 1). *drchrono Security Policy*. Retrieved from https://www.drchrono.com/security/

eMarketer. (2013, October 14). *Smartphone, Tablet Uptake Still Climbing in the US*. Retrieved from http://www.emarketer.com/Article/Smartphone-Tablet-Uptake-Still-Climbing-US/1010297

EyeVerify LLC. (2015). *Eyeprint ID*. Retrieved from https://itunes.apple.com/us/app/eyeprint-id/id864206217?mt=8`and https://play.google.com/store/apps/details?id=com.eyeverify.android_verify&hl=en

Farrell, M. B., & Wan, P. (2014, April 24). Hacker group Anonymous targets Children's Hospital. *The Boston Globe*. Retrieved from http://www.bostonglobe.com/business/2014/04/24/hacker-group-anonymous-targets-children-hospital-over-justina-pelletier-case/jSd3EE5VVHbSGTJdS5YrfM/story.html

Feng, T., Zhao, X., Carbunar, B., & Shi, W. (2013). Continuous Mobile Authentication Using Virtual Key Typing Biometrics. In *Proceedings of the 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications* (pp. 1547-1552). IEEE. 10.1109/TrustCom.2013.272

Finkle, J., & Hosenball, M. (2014, January). *Security Breach Hits More Prominent U.S. Retailers, Report Says*. Retrieved from http://www.huffingtonpost.com/2014/01/11/security-breach-more-retailers_n_4583200.html

Gadyatskaya, O., Massacci, F., & Zhauniarovich, Y. (2014). Security in the Firefox OS and Tizen Mobile Platforms. *IEEE Computer, 47*(6).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gartner. (2014, January 7). *Gartner Says Worldwide Traditional PC, Table, Ultramobile and Mobile Phone Shipments On Pace to Grow 7.6 Percent in 2014*. Retrieved from http://www.gartner.com/newsroom/id/2645115

Google, Inc. (2015). *Google Authenticator*. Retrieved from https://play.google.com/store/apps/details?id=com.google.android.apps.authenticator2&hl=en and https://itunes.apple.com/us/app/google-authenticator/id388497605?mt=8

Greenway Health. (2015). *PrimeSUITE*. Retrieved from http://www.greenwayhealth.com/solutions/primesuite/

Health I. T. (2015, March). *Hospital EHR Vendors*. Retrieved from http://dashboard.healthit.gov/quickstats/pages/FIG-Vendors-of-EHRs-to-Participating-Hospitals.php

HHS.gov. (2013). *Health Information Privacy*. Retrieved from http://www.hhs.gov/hipaa/

Hwang, S. S., Cho, S., & Park, S. (2009). Keystroke dynamics-based authentication for mobile devices. *Computers & Security*, *28*(1-2), 85–93. doi:10.1016/j.cose.2008.10.002

Information Security Media Group. (2015). *Healthcare Information Security Today – 2015 Survey Analysis: Evolving Threats and Health Info Security Efforts*. Retrieved from http://6dbf9d0f8046b8d5551a-7164cafcaac68bfd3318486ab257f999.r57.cf1.rackcdn.com//2015-healthcare-information-security-today-survey-pdf-5-h-60.pdf

iTELAGEN Healthcare IT Redefined. (2015). *iTELAGEN Healthcare IT Redefined*. Retrieved from http://www.greenwaysupport.com/home/

Jeggenstuen, J. (2013, December). *One In Every 5 People In The World Own A Smartphone, One in Every 17 Own A Tablet*. Retrieved from http://www.businessinsider.com/smartphone-and-tablet-penetration-2013-10

Jeong, H., & Choi, E. (2012). User Authentication using Profiling in Mobile Cloud Computing. In *Proceedings Of AASRI Conference on Power and Energy Systems* (Vol. 2, pp. 262-267). 10.1016/j.aasri.2012.09.044

Kareo. (2015, February 28). *Kareo EHR*. Retrieved from http://www.kareo.com/ehr/

Kelly, S. M. (2014, June 27). *In Google Fit vs. Apple HealthKit, Fitness Apps Stay Neutral*. Retrieved from http://mashable.com/2014/06/27/healthkit-google-fit-apps/#nf_r0U7flmqC

Li, P., Tian, J., Yang, X., Shi, P., & Zhang, Y. Y. (2009, June). Biometric Template Protection. *Journal of Software*, *20*(6), 1553–1573. doi:10.3724/SP.J.1001.2009.03528

Lomas, N. (2015, March). *Jolla Partners To Build Secure Version Of Its Sailfish Mobile OS*. Retrieved from http://techcrunch.com/2015/03/02/jolla-partners-to-build-secure-version-of-its-sailfish-mobile-os/

Maiorana, E., Campisi, P., González-Carballo, N., & Neri, A. (2011). Keystroke dynamics authentication for mobile phones. In *Proceedings of the 2011 ACM Symposium on Applied Computing, SAC 2011* (pp. 21-26). ACM.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

McCormack, M. (2014). *EHR Meaningful Use Market Share – 2014*. Retrieved from http://www.softwareadvice.com/resources/ehr-meaningful-use-market-share/

Medical, M. D. S. (2014). *mdsCONNECT – Mobile PrimeSUITE*. Retrieved from http://www.mdsmedicalsoftware.com/solutions/mdsconnect/

Mozilla. (2015). *Firefox OS*. Retrieved from https://www.mozilla.org/en-US/firefox/os/

O'Neill, P. H. (2014, March). *Beta testing for Tor's anonymous mobile OS begins this spring*. Retrieved from http://www.dailydot.com/technology/tor-anonymous-mobile-os-tails/

Online Tech. (2011, July 8). *What is HIPAA Compliance?* Retrieved from http://www.onlinetech.com/resources/references/what-is-hipaa-compliance

Ponemon Institute. (2009, April 6). *Ponemon Institute*. Retrieved from http://www.ponemon.org/

Ponemon Institute. (2015, May). *2015 Cost of Data Breach Study: Global Analysis*. Retrieved from http://public.dhe.ibm.com/common/ssi/ecm/se/en/sew03053wwen/SEW03053WWEN.PDF?

Practice Fusion. (2009). *Inside EHR Security*. Retrieved from http://www.practicefusion.com/whitepapers/ehr-and-system-security/

Practice Fusion. (2013). *Industry-leading EHR data security*. Retrieved from http://www.practicefusion.com/secure-ehr/

Qualcomm. (2015, August 4). *Future of security: Can we say good-bye to passwords?* Retrieved from https://www.qualcomm.com/news/onq/2015/08/04/future-security-can-we-say-good-bye-passwords

Real I. T. Experts. (2013). *Fitness Tracker*. Retrieved from https://play.google.com/store/apps/details?id=com.realitinc.fitnesstracker

Reisinger, D. (2014, May). Samsung eyes new biometric sensors for mobile devices. *Cnet*. Retrieved from http://www.cnet.com/news/samsung-eyes-biometric-sensors-for-future-mobile-devices/

Rindfleisch, T. C. (1997). Privacy, Information Technology, and Health Care. *Communications of the ACM*, *40*(8), 93–100. doi:10.1145/257874.257896

Rouse, M. (2015, July 21). *What is Health information technology (health IT or HIT)?* Retrieved from http://searchhealthit.techtarget.com/definition/Health-IT-information-technology

RT. (2014, October 22). *Mobile cyber-attacks surge as hackers chase banking info*. Retrieved from https://www.rt.com/news/198180-mobile-cyber-attacks-hackers/

Ruggiero, P., & Foote, J. (2011). *Cyber Threats to Mobile Phones.* US-CERT, United States Computer Emergency Reeiness Team, ©, Carnegie Mellon University. Retrieved from https://www.us-cert.gov/sites/default/files/publications/cyber_threats-to_mobile_phones.pdf

Sailfish. (2015, January 29). *Jolla Sailfish*. Retrieved from https://sailfishos.org/

Schaffer, K. B. (2015). Expanding Continuous Authentication with Mobile Devices. *Computer*, *48*(11), 92–95. doi:10.1109/MC.2015.333

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Siddiqui, Z., Abdullah, A. H., Khan, M. K., & Alghamdi, A. S. (2014, January). Smart Environment as a Service: Three Factor Cloud Based User Authentication for Telecare Medical Information System. *Journal of Medical Systems*, *38*(1), 1–14. doi:10.100710916-013-9997-5 PMID:24346931

Silent Circle. (2015, August 12). *Blackphone*. Retrieved from https://blackphone.ch/phone/

Stricker, P. (2010, July). *Understanding How Different Health Information Technology Applications Support Case Management*. Retrieved from http://www.cmsa.org/Individual/NewsEvents/HealthTechnologyArticles/tabid/533/Default.aspx

Sun, J., Zhang, R., Zhang, J., & Zhang, Y. (2014). *Touchin: Sightless two-factor authentication on multi-touch mobile devices*. Retrieved from http://arxiv.org/abs/1402.1216

Symantec. (2013). *VIP Access for Mobile*. Retrieved from http://m.vip.symantec.com/home.v

Tizen. (2014, May 31). *Tizen*. Retrieved from https://www.tizen.org/

Trojahn, M., & Ortmeier, F. (2013). Toward Mobile Authentication with Keystroke Dynamics on Mobile Phones and Tablets. In *Advanced Information Networking and Applications Workshops (WAINA*), *2013 27th International Conference* (pp. 697-702). Academic Press.

Ubuntu. (2015, February 9). *Ubuntu Touch*. Retrieved from http://www.ubuntu.com/phone

Ubuntu Developer. (2013). *AppArmor policy for click packages*. Retrieved from https://developer.ubuntu.com/en/start/platform/guides/app-confinement/

Ubuntu Developer. (2014). *Content Hub Guide*. Retrieved from https://developer.ubuntu.com/en/start/platform/guides/content-hub-guide/

U.S. Department of Education. (2010, April 7). *American Recovery and Reinvestment Act (ARRA)*. Retrieved from http://www.ed.gov/open/plan/recovery-gov

VASCO. (2014, November 3). *DIGIPASS for Apps: An all-in-one developer's toolkit for mobile*. Retrieved from https://www.vasco.com/products/developer_tools/digipass_for_apps/digipass_for_apps.aspx

VASCO. (2015). *VASCO*. Retrieved from https://www.vasco.com/

Whittaker, Z. (2013, December). Biometric smartphones to become mainstream in 2014. *ZDNet*. Retrieved from http://www.zdnet.com/article/biometric-smartphones-to-become-mainstream-in-2014-ericsson-says/

Yu, J., Wang, G., & Mu, Y. (2012). Provably Secure Single Sign-on Scheme in Distributed Systems and Networks. In *Proceedings of IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (pp. 271-278). IEEE. 10.1109/TrustCom.2012.228

Zaibak, O. (2014, May). Top 5 Voice. *Biometrics*, *2014*. Retrieved from http://www.voicetrust.com/blog/top-5-voice-biometrics-trends/

Zeltser. (2015, March 15). *Beyond Logins: Continuous and Seamless User Authentication*. Retrieved from https://zeltser.com/continuous-user-authentication/

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zhao, X., Feng, T., & Shi, W. (2013). *Continuous mobile authentication using a novel graphic touch gesture feature*. IEEE BTAS. doi:10.1109/BTAS.2013.6712747

ADDITIONAL READING

Althobaiti, M., & Mayhew, P. (2014). Security and usability of authenticating process of online banking: User experience study. *2014 International Carnahan Conference on Security Technology (ICCST)* (pp. 1-6). 10.1109/CCST.2014.6986978

Jaramillo, D., Newhook, R., Nguyen, V., & Chopra, M. (2015). Password-based mobile access, alternatives and experiences. *SoutheastCon*, *2015*, 1–8.

Lee, J., Jeong, Y., & Park, J. (2014). A Rhythm-Based Authentication Scheme for Smart Media Devices. *TheScientificWorldJournal*, *2014*, 1–9. PMID:25110743

Li, F., Clarke, N., Papadaki, M., & Dowland, P. (2014). Active authentication for mobile devices utilising behaviour profiling. *International Journal of Information Security*, *13*(3), 229–244. doi:10.100710207-013-0209-6

Liang, H., Fleming, C., & Wang, W. (2014). User Authentication Interfaces in Mobile Devices: Some Design Considerations. *2014 IEEE 17th International Conference on Computational Science and Engineering* (pp. 754-757).

Meng, W., Wong, D., Furnell, S., & Zhou, J. (2014). *Surveying the Development of Biometric User Authentication on Mobile Phones* (pp. 1268–1293). IEEE Communications Surveys & Tutorials IEEE.

Saevanee, H., Clarke, N., Furnell, S., & Biscione, V. (2015). Continuous user authentication using multi-modal biometrics. *Computers & Security*, *53*, 234–246. doi:10.1016/j.cose.2015.06.001

Skracic, K., Pale, P., & Jeren, B. (2013). Knowledge based authentication requirements. *2013 36th International Convention on Information Communication Technology Electronics Microelectronics (MIPRO)* (pp. 1116-1120).

Technologies, C. A. (2014, September). *Healthcare Security Solutions: Protecting your Organization, Patients, and Information*. Retrieved from http://www.ca.com/us/~/media/Files/SolutionBriefs/healthcare-security-solutions.pdf

Trewin, S., Swart, C., Koved, L., Martino, J., Singh, K., & Ben-David, S. (2012). Biometric authentication on a mobile device: A Study of User Effort, Error and Task Disruption. *Proceedings of the 28th Annual Computer Security Applications Conference on - ACSAC '12* (pp. 159-168). 10.1145/2420950.2420976

Ziminski, T., Demurjian, S., & Sanzi, E. (2016). Toward Integrating Healthcare Data and Systems: A Study of Architectural Alternatives. In T. Ivamu & D. Tatnall (Eds.), *Maximizing Healthcare Delivery and Management through Technology Integration* (pp. 270–304). doi:10.4018/978-1-4666-9446-0.ch016

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zirjawi, N., Kurtanovic, Z., & Maalej, W. (2015). A survey about user requirements for biometric authentication on smartphones. *2015 IEEE 2nd Workshop on Evolving Security and Privacy Requirements Engineering (ESPRE)* (pp. 1-6). doi: 10.1109/ESPRE.2015.7330160

KEY TERMS AND DEFINITIONS

Biometrics: Technologies utilized with the purpose of analyzing human body characteristics (e.g., fingerprints, voice patterns, facial patterns).

Chirography: New biometric method, which is based on obtaining what a user writes on a multi-touch screen with his/her finger.

Continuous User Authentication: Authentication method that continuously authenticates a user through the means of some of his/her biometric traits.

Health Information Technology (HIT) Systems: Electronic systems that are utilized with the expectation that they will lower costs, improve efficiency, and reduce errors, in addition to providing better care and services for consumers in the healthcare domain.

Inherence Factor: User authentication method based on things a user is, which can include traits of the user such as fingerprints and eye retina.

Keystroke Dynamics-Based Authentication (KDA): Behavioral biometrics approach used to analyze the way that a person types text on a keyboard/touchscreen to verify a user's credentials.

Knowledge Factor: User authentication method based on things a user knows which can include having knowledge of a phrase/pattern utilized to unlock a device.

Mobile Computing: Devices that allow people to access software/data regardless of their location.

Multi-Factor Authentication (MFA): Method that uses two or more authentication factors (knowledge factors, possession factors, and inherence factors) in order to verify if a person is who he/she says he/she is in a place, system, etc.

Possession Factor: User authentication method based on items that a user has with him/her, which can include hardware device such as tokens.

Risk-Based Authentication (RBA): Method that provides different levels of authentication in a system depending on the confidentiality of the data the user is trying to have access to, and on a risk score threshold.

User Authentication: Process done to verify if a person has access to a certain place, system, application, etc.

User Profiling: Method that consists of discovering patterns in large quantities to establish if the user is who he/she says he/she is by tracking the way that he/she utilizes his/her mobile device, and deviations of behavior may indicate a problem.

This research was previously published in Innovative Solutions for Access Control Management edited by Ahmad Kamran Malik, Adeel Anjum, and Basit Raza; pages 160-196; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 63
Perceptions of Play:
Using Play-Doh to Enhance the Student Experience in Bioscience Higher Education

Gemma Lace-Costigan
University of Salford, UK

ABSTRACT

Playful and kinaesthetic learning approaches are used in numerous early years (birth to 5 years old) learning environments, however studies in HE STEM disciplines are uncommon. This study aimed to explore the use of Play-Doh in an undergraduate anatomy module as a method of enhancing engagement. 63 students attended the 'kinaesthetic play' lecture, where students worked in teams to make a variety of epithelial cell types using Play-Doh. Before and after the activity, students were asked to 'choose one word to describe how you feel'. Before the activity, 48.3% of responses were negative (E.g. confused, worried, childish). However, after the activity not a single negative response was recorded. 98% of students reported that they enjoyed the activity and 84% reported that the activity increased their understanding. This data suggests a utility for kinaesthetic playful practice in STEM teaching. Overcoming initial student perceptions towards alternative teaching practices is a challenge to be considered during session design.

INTRODUCTION

The concept of playful learning is well researched and has formed the foundation of early years educational teaching methods (Piaget, 1999, Kangas, 2010, Broadhead et al, 2010). Engagement of adult learners through playful practice has been demonstrated in a few studies (Rice, 2009; Nerantzi et al, 2015), however there is relatively little research into the role and potential for playful learning methods in higher education (HE), particularly in STEM subject teaching including biological and biomedical science. Defining 'play' for an adult learner is difficult but in 1999, Henricks proposed that play could be considered an experience that has intrinsic (rather than extrinsic) motives, so that someone would choose to participate in that activity for the experience of participation alone and that learning is not the primary driver of participation. Henricks also postulated that the process of play was more impor-

DOI: 10.4018/978-1-7998-1204-3.ch063

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tant than the end result and play should involve some degree of active engagement. When students are completely captivated by a playful activity, Csikszentmihalyi (2013) proposed a state of mind called 'flow' where deeper learning can be achieved in both adults and children. Incorporating playful learning activities in adult learning environments therefore has the potential to increase learning whilst enhancing student experiences and satisfaction which has been reported to occur as a consequence of 'flow' (Csikszentmihalyi, 2013).

Learning models that incorporate kinaesthetic elements such as the VARK (visual, aural, read/write, kinaesthetic) model developed by Neil Fleming, have been well described across numerous subject disciplines (Fleming & Mills, 1992). VARK learning style assessments are still used in HE in an attempt to enable learners with the information that will allow them to make informed choices about the way they select learning materials, study and prepare for exams. Within the biological and biomedical science disciplines, students are often encouraged to contextualise their learning through kinaesthetic means such as active participation in laboratory practical experiments or anatomy dissections. 'Learning through doing' or experiential learning (described by Kolb, 2014) is therefore a common practice within STEM teaching/learning. However, the kinaesthetic learning activities used often lack a creative dimension and often involve following step by step protocols to achieve a desired end result. There are limited opportunities for creative or explorative learning within some of the common kinaesthetic activities associated with science teaching and this may restrict the amount of deeper learning that can occur. There are countless texts describing the importance of providing opportunities for creativity within learning environments (Beghetto, 2015; Cropley, 2001; Mayer, 1989). However, it is unclear if these kinds of activities would be enjoyable or beneficial to HE STEM students, particularly those studying on biological or biomedical science programmes.

The field of cognitive psychology has generated a multitude of research describing how to increase engagement and promote learning. It is known that memories are enhanced during emotional arousal (Dolcos et al, 2004) and for a long time neurobiologists have researched the cellular basis of 'emotional learning' (LeDoux, 1992). Understanding that a student is more likely to remember a specific lecture (and associated content) if an emotional response is instigated during the lecture is something that can be exploited by academics when designing teaching sessions. Historically, individual students may have been targeted with questions during traditional lectures, or asked to stand in front of peers and take part in a demonstration which would most likely trigger a fear or an 'adrenaline' response in the participating student. The activation of a fear response may induce avoidance behaviours (LeDoux, 1992) which may lead to decreased lecture attendance. Whilst this method may have been successful in promoting increased attention in lectures (and memory of them afterward), this may not always have led to positive student experiences. With mental health issues such as anxiety on the increase in undergraduate student cohorts (Williams et al, 2015), it may be time to try and promote more positive emotional responses in lectures.

There are a number of challenges posed by teaching STEM subjects, for example, having to teach large cohort numbers. The need to teach large numbers of students at any one time may restrict the kind of learning environments available for use, for example, teaching may be restricted to the use of large lecture theatres with static furniture. Issues associated with 'crowd control' may deter academics from trying more adventurous and novel teaching styles. Teaching staff may also be required to cover a significant amount of complex course content in a relatively short period of time and the requirements of accrediting bodies may mean the curriculum lacks flexibility. Many STEM students may begin to feel overwhelmed by the volume and complexity of course material that is covered in a single teaching session. Additionally, research has reported that 59% of students find lectures boring half the time and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Play-Doh has a distinct and recognisable appearance, texture and smell

30% find lectures boring most of the time, with the use of PowerPoint slides being largely attributed to this effect (Mann & Robinson, 2009). In order to facilitate active learning and a deeper understanding of the more difficult (or tedious) course content it is vital that educators explore more novel methods of engaging students within the traditional lecture theatre environment whilst assessing the success of these more alternative approaches to teaching.

As described previously, there is evidence that both playful learning practice and kinaesthetic learning models have been successful in promoting engagement and learning in HE. Research has also demonstrated that learning is more likely to take place during a state of heightened emotional arousal (Dolcos et al, 2004) and so designing learning activities that stimulate a sense of happiness, creativity and nostalgia may be used to enhance learning of some of the least interesting/stimulating areas of a STEM programme. Play-Doh is a type of popular modelling clay that has been on the market since the 1950's and is sold in approximately 75 different countries worldwide (Walsh, 2005). For decades, children have played with this product in their own homes and also have been exposed to it in schools and other learning environments. Play-Doh has a distinct appearance, smell and texture that would be recognisable by most individuals born after the 1950's and it could be predicted that the use of this medium could trigger feelings of happiness and nostalgia whilst adding a creative and tactile element to a learning activity (Figure 1).

This study aimed to explore and evaluate the use of Play-Doh in a 'kinaesthetic play' lecture in an attempt to increase engagement and student enjoyment during a two hour histology lecture focusing on epithelial tissue.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

METHODS

Study Sample

This study was performed at the University of Salford with L4 students enrolled on the BSc (Hons) Biomedical Science (IBMS) programme which is offered by the School of Environment and Life Sciences. This programme is accredited by the Institute of Biomedical Science (IBMS) and so much of the curriculum content is inflexible to ensure that graduates have covered the key learning material for graduate biomedical science practice. The most recent National Student Survey reported 90% overall student satisfaction with this programme and that 91% of students were employed or in further study within six months of graduating (Which? University, 2016). Common careers for biomedical science graduates have included employment as NHS biomedical scientists, research scientists, industry scientists and employment as medical writers. Other graduates have gone on to study for MRes, MSc, PhD, Medicine, Dentistry and Teaching qualifications.

121 students were enrolled on the L4 Human Anatomy module which is a lecture based module delivered over two semesters. 2.3% of these students were international students, 60.3% were female and 18% were classed as mature students. The group was ethnically diverse and included 29% Asian/ Asian British Pakistani students, 23.3% White British students, 12.2% Black/Black British African student and 7% being of Asian/Asian British Indian origin.

Kinaesthetic Play Lecture Design

In the first semester of the L4 Human Anatomy module the teaching and learning focus was on gross anatomy and was delivered by a single member of staff via traditional lecture methods in a large lecture theatre. In the second semester, the focus shifted to tissue histology and was again delivered by a single (different) member of staff, Dr Gemma Lace-Costigan (GLC) in the same lecture theatre. The lecture theatre capacity was sufficient to seat 130 students and was divided into two equal seating blocks (separated by stairs with additional stairs at the end of each block), with each block being organised into rows of 8-10 seats. Entry points were located at both the front and the back of the room.

In the first lecture delivered by GLC, students were advised that different teaching methods would be used in an attempt to enhance learning and promote engagement. As per the University of Salford academic regulations (University of Salford, 2016), lecture notes were made available at least twenty-four hours before the lecture, however students were informed that details regarding some of the planned learning activities would be removed to ensure an element of surprise and originality to the sessions.

The kinaesthetic play lecture was the fourth lecture to take place in the series of 8 lectures delivered. Out of the 121 students enrolled on the module, 63 students attended the kinaesthetic play lecture. Before the start of the lecture, each seating space had a single sticky note placed on the desk which was to be used to collect the study feedback. At the start of the session students were informed that the content delivered in the first half of the session would be used to allow all students to participate in a competition later on in the lecture. Students were then given a traditional lecture (via PowerPoint) describing the role, types and features of different types of epithelial tissue. Once the subject material had been delivered, students were informed that the competition was about to start. The students were presented with the following information on a PowerPoint slide: ‘Working with the students in your row, you have 10 mins to create each of the following cell types; simple squamous epithelia, simple ciliated columnar epithelia,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

simple cuboidal epithelia, stratified cuboidal epithelia. Label in detail. Don't forget the 'general' features of an epithelial cell'. Students were given the opportunity to ask any questions about the activity and then before the activity started a new slide was displayed which directed the students to use their sticky notes to anonymously complete the following task; Choose ONE word to describe how you feel about taking part in this activity. Once this was completed the competition instructions slide was projected once again and students were asked a second time about any activity queries and a small pot of Play-Doh was given to each student. Once all queries were answered a timer was started and the activity began. During the activity, the academic in charge (GLC) walked around the room, encouraging participation and asking/answering student questions. All students present took part in the activity. After 10 minutes, all students brought their entries down to the front of the room for formative assessment. The work was primarily assessed by GLC but students were also invited to leave feedback for their peers with respect to the accuracy of the models and labelling. It was reinforced that models would not be assessed based on their artistic merit but instead on evidence of understanding the structure of the various epithelial cells. The entries were discussed as a group and a small prize (a choice of chocolate or stationary) was awarded to the team that produced the best models. The students were then asked to complete the following task; choose ONE word to describe how you feel AFTER taking part in the activity. Students were also asked to rate (1-strongly agree, 2-agree, 3- undecided, 4- disagree or 5- strongly disagree) the following statements; 1) I enjoyed that activity and 2) That activity increased my understanding of epithelia. The feedback was collected and the lecture then continued using traditional PowerPoint delivery.

Data Analysis

Of the 63 participants, 58 fully and correctly completed feedback sticky notes. 5 sets of feedback were not utilised in the study due to missing ratings or use of full sentence responses rather than a single word response.

A basic language analysis was performed by inputting the responses into WordClouds.com to assess the frequency of responses for each question and to generate a visual summary of responses. This site generates word clouds allowing the visualisation of text inputs so that the most frequently occurring words appear in larger font, whereas the less frequently occurring words appear in a smaller font. Each response was categorised as being optimistic/positive, neutral/ambiguous in meaning or pessimistic/negative.

A register was taken at the end of the session and this was used to explore the relationship between kinaesthetic play lecture attendance and overall module exam performance via an unpaired T-test.

Results

All attending students participated in the learning activity (100% engagement) with each individual student producing one of the four epithelial cell types required. In order to submit a competition 'entry,' each team had to submit all four epithelial cell types. When there were more than four members in each team, additional cells were made to ensure all individuals could create something as demonstrated in Figure 2.

After being presented with the learning activity students were ask to 'Choose ONE word to describe how you feel about taking part in this activity' (Figure 3A). After participating in the learning activity students were then asked to 'choose ONE word to describe how you feel AFTER taking part in the activity' (Figure 3B). Figure 3A highlights that before the activity, the most common responses were 'excited,' 'confused' and 'shocked'. However, after the activity, the most common responses were 'happy,'

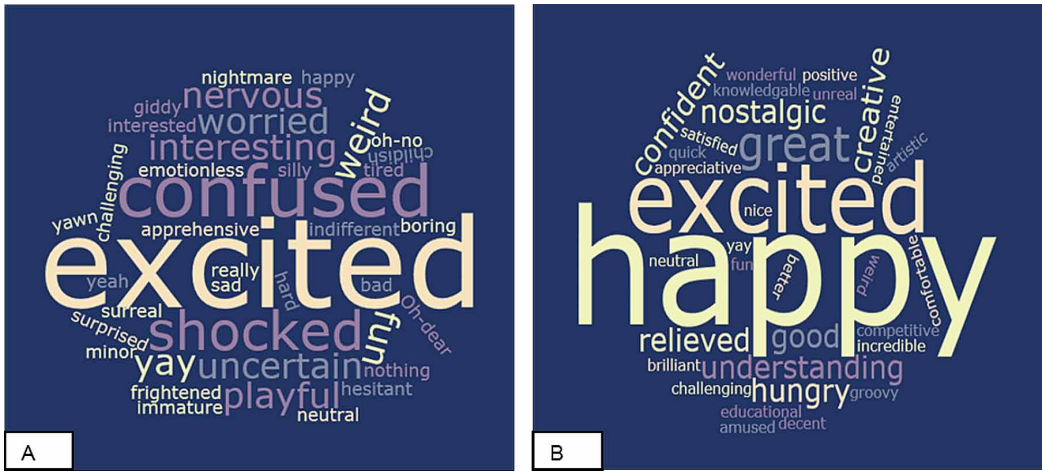
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The kinaesthetic play lecture activity

'excited' and 'great'. The full range of student responses can be visualised in Figure 3, where the most commonly reported words appear in a larger font size within the word cloud and the least commonly reported words being reported appearing in a smaller font size.

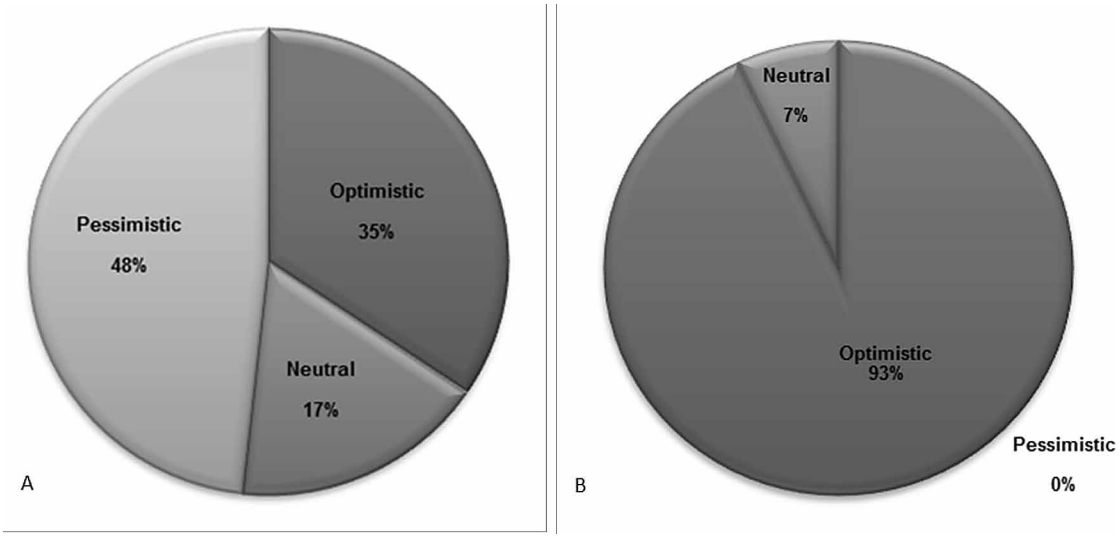
When asked to 'choose ONE word to describe how you feel about taking part in this activity' before the activity took place, 48% of responses used language that was negative/pessimistic, 35% that was positive/optimistic and 1% of responses were neutral or ambiguous in meaning (Figure 4A). The most popular three negative/pessimistic words were 'confused,' 'shocked' and 'worried' and a number of students also reported feeling 'nervous,' 'apprehensive,' 'frightened' and also selected words such as 'nightmare,' 'immature,' 'childish' and 'boring'. The most common three optimistic words were 'excited,' 'interesting' and 'playful' and other responses included 'fun,' 'yay,' 'giddy' and 'happy'. The most frequently selected neutral/ambiguous words were 'weird,' 'indifferent' and 'emotionless'. Overall, the top three responses were 'excited' which was interpreted as being a positive/optimistic response, and 'confused' and 'shocked' which were deemed negative/pessimistic responses.

Figure 3. Perceptions of play A) before the activity B) after the activity (Word cloud created at word-clouds.com)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Language analysis of responses A) before the activity and B) after the activity

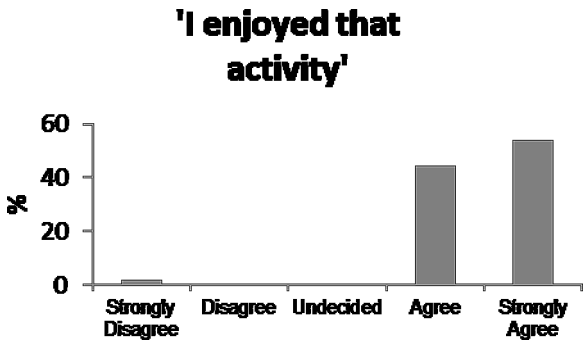


When asked to 'choose ONE word to describe how you feel AFTER taking part in this activity,' 0% of responses were negative/pessimistic, 93% were positive/optimistic and the remaining 7% responses were neutral or ambiguous in meaning (Figure 4B). The most commonly selected positive/optimistic responses were 'happy,' 'excited' and 'great' and other responses included 'understanding,' 'nostalgic,' 'confident' and 'creative'. The most frequent neutral/ambiguous responses were 'hungry,' 'neutral' and 'weird'.

When participants were asked to consider the statement 'I enjoyed that activity,' 54% of participants reported to 'strongly agree,' 44.4% reported to 'agree,' 1.6% reported to 'strongly disagree' and no students selected to 'disagree' or were 'undecided' (Figure 5). Overall, 98.4% students therefore reported enjoying the activity.

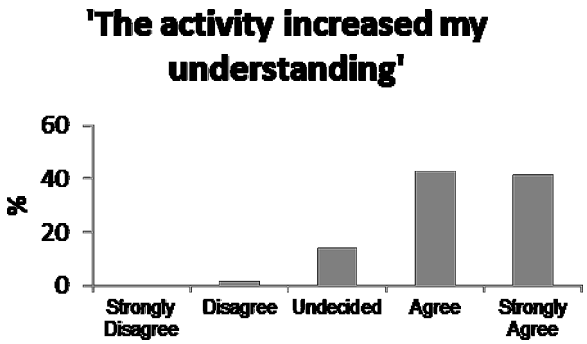
When participants were asked to consider the statement 'The activity increased my understanding of epithelia,' 41.3% students reported that they could 'strongly agree,' 42.9% reported to 'agree,' 14.3% were 'undecided, 1.6% reported to disagree and no students chose to 'strongly disagree' with the statement. Therefore 84.2% of students felt the activity increased their understanding of epithelia (Figure 6).

Figure 5. Student ratings of the statement 'I enjoyed that activity'



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Student ratings of the statement 'The activity increased my understanding of epithelia'



Out of the 121 students enrolled on the module, 63 students attended the playful lecture. The group of students that attended the session achieved a mean exam result of 70.1% whereas the student group which did not attend the lecture attained a significantly lower exam mark of 60.1% ($p<0.05$). Access to the full lecture series attendance records were not available and so marks could not be adjusted to compensate for regular lecture non-attendance.

CONCLUSION

Designing engaging and interactive lecture sessions that enhance deep learning is a challenge taken on by academics across different discipline areas worldwide. Kinaesthetic learning activities and playful learning practice has been shown to be successful in a number of disciplines, yet these methods are poorly explored in HE STEM subjects. Teaching to large numbers of students within the constraints of a traditional and static lecture theatre can be intimidating at times, and educators may be cautious with respect to introducing kinaesthetic and playful learning activities due to issues associated with 'crowd control'. Modern day Universities strive to increase student satisfaction as well as exam performance and some academics may feel there is a risk associated with deviating from the normal modes of educational practice. More research is needed to explore the potential use of alternative teaching methods with larger student groups within traditional lecture hall environments in HE STEM disciplines.

Academics whom are keen to appease their students may be concerned that the incorporation of kinaesthetic and playful learning activities may be considered patronising or childish by adult learners. In this study, Play-Doh was used as a learning material and this particular medium is a popular toy that is also used in early years' education. After being briefed about the learning activity, the initial responses from students regarding taking part in the kinaesthetic play activity were mostly negative/pessimistic (48.3% of responses), with only 34.5% of students giving a positive/optimistic response. Many of the responses given were anxious responses (for example 'confused,' 'shocked,' 'uncertain,' 'worried,' 'nervous,' 'apprehensive,' 'frightened') indicating students were being pushed out of their learning comfort zones. The Learning Zone Model of learning (Senninger, 2000) describes how students can be encouraged to learn by venturing out of their comfort zone but without being pushed into their 'panic' zone. It would be interesting to note the impact on the responses if the students were given advanced warning of the class activity in future sessions, thus removing the element of surprise and not pushing students into their

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

'panic' zone. However, by doing this you would introduce the risk of students harbouring any negative preconceptions simply not turning up to the session. Though the majority of teaching on the BSc (Hons) Biomedical Science Programme at the University of Salford is centred around traditional lectures, lab classes and tutorials, less traditional approaches such as flipped-learning, workshops, poster/oral presentation sessions, problem based learning, seminars etc are also increasing being incorporated within the programme. However, the majority of these alternative teaching practices are utilised at L5 and L6 where the class sizes are smaller and the subject material is more specialised. Increasing the variety of teaching styles and learning activities right from the start of L4 may be useful to modify student expectations so that alternative learning activities are not perceived as threatening, intimidating and without pedagogical benefit. Encouraging and inviting students to step out of their learning comfort zones whilst being mindful that certain activities may hinder learning due to negative emotional reactions is likely to enhance both learning and student learning experiences (Senninger, 2000).

A number of the responses collected before the learning activity indicated that students perceived the activity to be childish or unsuitable at the HE level. Responses such as 'immature,' 'childish' and 'silly' reinforce the importance of explaining the significance of the activity with respect to its potential to enhance learning. If a student does not consider a learning activity meaningful and worthy of their time and effort, they are less likely to fully engage and commit to the interaction (Fredricks et al, 2004). Making students aware of the success of other adult associated playful learning initiatives such as 'Lego Serious Play' (James, 2015; Nerantzi et al, 2015), games-based learning approaches (Whitton & Moseley, 2012) and the use of augmented reality to enhance scientific laboratory skills (Akçayır et al, 2016) may be effective in demonstrating the significance of such playful approaches. By carefully explaining the underlying pedagogical principles associated with playful/kinaesthetic learning activities and how this relates to achieving intended learning outcomes, one might expect to reduce some of the initial negative/pessimistic perceptions evoked.

It is extremely interesting to note that while 48.3% of the responses were negative/pessimistic before the activity, this value dropped to 0% after the activity with 93% of students giving positive/optimistic responses. Many of the responses indicated a level of enjoyment (for example, responses included feeling 'happy,' 'excited,' 'great,' 'incredible,' 'positive,' 'brilliant,' or the word 'fun') and a number of students reported the feeling of nostalgia. Creativity was also a theme amongst responses (for example students reported feeling 'creative' and 'artistic') and a significant number of responses indicated that the students felt they had learned from the activity (responses included the words 'understanding,' 'knowledgeable' and 'educational'). The lack of negative/pessimistic responses after the learning event demonstrates that once the learners had experienced the learning activity first hand, those negative feelings were extinguished. The source of those original negative feelings may have been due to the association between Play-Doh, childhood memories and early years' educational experiences and a misalignment with their expectations of adult learning methods in University. The kinaesthetic play lecture would have been the first and only exposure to a creative and tactile learning activity whilst in a large group teaching session in HE and some level of scepticism is not unexpected. Students come into HE with certain preconceptions of how, when and where learning will take place. It would have been interesting to compare enjoyment levels before the introduction of the playful learning activity to how students perceived the lecture with an incorporated kinaesthetic playful element. It would also be interesting to follow up with those students in the future to see how their perceptions change with time. It is only natural for a student to relate back to their own personal experiences of education or media portrayals of University life and in a culturally and ethnically diverse population such as cohort being described in this study, these preconceptions may

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

vary significantly. These findings again reinforce the significance of increasing student awareness of the success of other adult learner associated playful learning practices (such as 'Lego Serious Play') and the scholarship associated with these methods. By seeing these learning activities as part of a research informed approach, students in HE may have less negative preconceptions about participation. It is an extremely significant finding that despite almost half of the cohort originally having negative thoughts and feelings associated with the kinaesthetic play activity, after the activity the majority of the group gave positive feedback. It could therefore be postulated that increased exposure to alternative teaching styles (and understanding of the underlying pedagogies) may be more likely broaden the minds of learners in HE so that similar activities are not viewed so negatively when presented in the future. It would have been useful to conduct a series of pre-activity and post-activity interviews to allow a more detailed exploration of student perceptions of kinaesthetic playful practice in bioscience HE. However, given the large cohort size, this would have proved logistically difficult and so a follow up study targeting smaller group sessions in bioscience HE might be useful in dissecting out what underlies some of the negative preconceptions of playful practice.

This study aimed to explore whether HE STEM students would appreciate and enjoy a more artistic and tactile learning activity, despite the subject area conventionally not being associated with creative learning approaches. 98.4% of the students that participated in the kinaesthetic play lecture reported that they enjoyed the activity which demonstrates more creative learning activities should not be overlooked simply because they do not traditionally sit within STEM teaching methods. The increasingly collaborative nature of academia means that science and art fusion projects are on the increase therefore enabling students to think about science creatively is likely to increase employability. Encouraging students to express scientific understanding through artistic means may help consolidate their learning, since many biological/biomedical students may in fact learn through kinaesthetic learning methods (Kolb, 1994). There exists a catalogue of historical research describing different learning styles and preferences (Honey & Mumford; 1982, Kolb, 1981; Myers-Briggs, 1980; Schmeck, 1998) however more modern approaches celebrate the complex and individual nature of learning and how learning preferences may be dynamic rather than static. It is therefore important for educators to provide a variety of learning activities that so that students can explore which learning methodologies work best for them and their individual learning requirements at any given stage of their learning journey.

Using Play-Doh as a learning tool served to inject some creativity within a traditional lecture based learning environment, allowing students to individually participate in the learning activity whilst also working as part of a team. A wide variety of literature is available to support that many students 'learn through doing' (Kolb, 2014) and since each student would be individually producing a Play-Doh model, each student would therefore be actively involved in the session. By immersing themselves in the tactile and sensory task of making different epithelial cells, the student was creating their own, personal experience and memory of the kinaesthetic play lecture which would facilitate their learning through active participation. Additionally, each individual formed part of a team so even though parts of the activity were being performed in solitude, there was a collaborative slant on the activity. When students work well with others and experience a sense of connection whilst undertaking a learning task, engagement has been reported to increase (Wentzel, 2009; Deci & Ryan, 2000). Introducing a competitive element to the activity (via offering a small reward for the most accurate Play-Doh creation) increased the motivation of the student to participate through the desire to win and/or receive an award. The use of competitive activities to enhance motivation and learning has been described in numerous publications (Brophy, 2013; Weiner, 1990). Other studies have reported how cooperative learning techniques (which involve

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the use of activities where students work together on tasks to receive rewards or recognition) have been used successfully in primary and secondary school education (Slevin, 1980). Games participation is a classic example of how reward and 'winning' can be used to motivate engagement and newer pedagogical research approaches provide evidence to support the use of games to enhance teaching and learning (Whitton, 2011; Moseley & Whitton, 2012). It would be interesting to note if the results would have been different if the competitive element of the activity was removed. The kinaesthetic play activity used in this session combined using active participation, collaborative learning with a competitive element, which ultimately resulted in enhanced positive/optimistic feedback after the activity. Though team activities can be logistically difficult to manage in a traditional lecture hall environment, this study demonstrates that they are a worthwhile effort and useful in enhancing engagement, promoting student satisfaction and enhancing learner experiences.

A large number of responses after the activity were highly emotive (for example, students reported feeling 'happy,' 'excited,' 'nostalgic,' 'satisfied' or responded with the word 'fun') and previous research has demonstrated that in situations of heightened emotional arousal, learning and memory formation becomes more effective (Dolcos, 2004; LeDoux, 1992). The students who attended the lecture attained significantly higher marks in their exams compared to those who did not attend the lecture however these results should be interpreted with caution. It must be considered that only ~10% of the exam questions were associated with the kinaesthetic play lecture content and the non-attending students may have generally had poor attendance to the full lecture series. In future sessions, a before and after activity test could be delivered using a platform such as Kahoot. This would enable an exploration of how these kinds of tactile and playful activities may enhance understanding of the course content in a traditional lecture environment. It should be recollected that the main aim of this study was to assess the utility of using Play-Doh based kinaesthetic playful activities to enhance student engagement and enjoyment rather than to increase student exam performance, though it is known that a relationship between engagement and performance exists. Enhancing student learning should be the primary objective of a teaching activity and 84.2% of students felt the activity enhanced their understanding of the subject area and the other study findings support the notion that they had fun whilst learning.

Since 100% of the students present in the session actively participated in the learning activity, with 98.6% of students reporting that that they enjoyed the activity and 84.2% reporting they felt the activity increased their understanding of the subject, it can be concluded that kinaesthetic playful learning activities using Play-Doh can be used in increased student engagement, enjoyment and understanding. Engaging large numbers of HE students in creative, playful and kinaesthetic learning activities can be logistically challenging but ultimately beneficial to the learner. This research demonstrates that despite having some initial negative perceptions about kinaesthetic learning practice, students ultimately enjoy these alternatives to traditional lecture based delivery styles and report that they feel their understanding of the subject area has been enhanced. Kinaesthetic playful practice therefore has the potential to be used at different levels of education across a variety of different discipline areas, including within HE STEM teaching.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Akçayır, M., Akçayır, G., Pektaş, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior*, *57*, 334–342. doi:10.1016/j.chb.2015.12.054

Barbe, W. B., Milone, M. N., & Swassing, R. H. (1988). *Teaching through modality strengths: Concepts and practices*. Zaner-Bloser.

Beghetto, R. A. (2016). Creative learning: A fresh look. *Journal of Cognitive Education and Psychology*, *15*(1), 6–23. doi:10.1891/1945-8959.15.1.6

Broadhead, P., Howard, J., & Wood, E. (Eds.). (2010). *Play and learning in the early years: From research to practice*. Sage.

Brophy, J. E. (2013). *Motivating students to learn*. Routledge.

Cropley, A. J. (2001). *Creativity in education & learning: A guide for teachers and educators*. Psychology Press.

Csikszentmihalyi, M. (2013). *Flow: The psychology of happiness*. Random House.

Deci, E. L., & Ryan, R. M. (2000, October). Deci, E. L., & Ryan, R. M. (2000). The what and why of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. doi:10.1207/S15327965PLI1104_01

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*(5), 855–863. doi:10.1016/S0896-6273(04)00289-2 PMID:15182723

FlemingN. D.MillsC. (1992) VARK. Retrieved from www.vark-learn.com

Fleming, N. D., & Mills, C. (1992). Not another inventory, rather a catalyst for reflection.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. doi:10.3102/00346543074001059

Henricks, T. S. (1999). *Play as ascending meaning: Implications of a general model of play. In Play contexts revisited* (pp. 257–277). Stamford: Ablex Publishing Group.

Honey, P., & Mumford, A. (1982). *Manual of Learning Styles*. London: P. Honey.

Kangas, M. (2010). Creative and playful learning: Learning through game co-creation and games in a playful learning environment. *Thinking Skills and Creativity*, *5*(1), 1–15. doi:10.1016/j.tsc.2009.11.001

Kolb, D. A. (1981). Learning styles and disciplinary differences. *The modern American college*, *1*, 232-255.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

LeDoux, J. E. (1992). Brain mechanisms of emotion and emotional learning. *Current Opinion in Neurobiology*, *2*(2), 191–197. doi:10.1016/0959-4388(92)90011-9 PMID:1638153

Mann, S., & Robinson, A. (2009). Boredom in the lecture theatre: An investigation into the contributors, moderators and outcomes of boredom amongst university students. *British Educational Research Journal*, *35*(2), 243–258. doi:10.1080/01411920802042911

Miller, M. D. (2011). What college teachers should know about memory: A perspective from cognitive psychology. *College Teaching*, *59*(3), 117–122. doi:10.1080/87567555.2011.580636

Myers-Briggs, I. (1980). *Gifts Differing*. Palo Alto, CA: Consulting Psychologists Press.

Nerantzi, C., Moravej, H., & Johnson, F. (2015). Play Brings Openness or Using a Creative Approach to Evaluate an Undergraduate Unit and Move Forward Together. *Journal of Perspectives in Applied Academic Practice*, *3*(2). doi:10.14297/jpaap.v3i2.141

PiagetJ. P. (1999). De la pedagogía.

Rice, L. (2009). Playful learning. *Journal for Education in the Built Environment*, *4*(2), 94–108. doi:10.11120/jebe.2009.04020094

Schmeck, R. R. (1998). *Learning Strategies and Learning Styles*. New York: Plenum Press.

Senninger, T. (2000). *Abenteuer leiten-in Abenteuern lernen: Methodenset zur Planung und Leitung kooperativer Lerngemeinschaften für Training und Teamentwicklung in Schule, Jugendarbeit und Betrieb*. Ökotopia Verlag.

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, *50*(2), 315–342. doi:10.3102/00346543050002315 PMID:6698102

University of Salford. (2016), Academic Handbook. Retrieved from http://www.governance.salford.ac.uk/page/academic_handbook

Walsh, T. (2005). Play-doh. In Timeless toys: Classic toys and the playmakers who created them.

Weiner, B. (1990). History of motivational research in education. *Journal of Educational Psychology*, *82*(4), 616–622. doi:10.1037/0022-0663.82.4.616

Wentzel, K. R. (2009). Peers and academic functioning at school. In K. Rubin, W. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups. Social, emotional, and personality development in context* (pp. 531–547). New York, NY: Guilford Press.

Which? University (2016). Retrieved from http://university.which.co.uk/university-of-salford-s03/biomedical-science-3-years-8500-b900

Whitton, N. (2011). Encouraging engagement in game-based learning. *International Journal of Game-Based Learning*, *1*(1), 75–84. doi:10.4018/ijgbl.2011010106

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Whitton, N., & Moseley, A. (2012). *Using games to enhance learning and teaching: a beginner's guide*. Routledge.

Williams, M., Coare, P., Marvell, R., Pollard, E., Houghton, A., & Anderson, J. (2015). *Understanding provision for students with mental health problems and intensive support needs*. Brighton: Institute for Employment Studies.

This research was previously published in the International Journal of Game-Based Learning (IJGBL), 7(3); edited by Patrick Felicia; pages 26-37; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 64
Pre-Implementation Case Studies Evaluating Workflow and Informatics Challenges in Private Primary Care Clinics for Electronic Medical Record Implementation

Calvin Or
The University of Hong Kong, China

ABSTRACT

Despite the potential benefits of electronic medical records (EMRs), their implementations have failed in a number of instances because of misfit between the technology design, work practices, workflows, and clinicians' needs. To promote effective design and implementation of an EMR system, the present pre-implementation study modeled the clinical workflow processes in private primary health care clinics and identified the associated gaps, weaknesses, and informatics challenges in the processes. Systematic field observations were performed to collect workflow data. Forty-three health care professionals (16 physicians and 27 clinical assistants) of fifteen private primary health care clinics were studied. The results, presented in a workflow diagram, demonstrate step-by-step details of the clinical work processes, their gaps and weaknesses, and possible improvement opportunities offered by the application of EMR. The workflow model provides important implications that may be applicable to other health care settings that plan to implement EMRs.

DOI: 10.4018/978-1-7998-1204-3.ch064

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

An initiative that focuses on implementation of an electronic medical record (EMR) sharing system in both public and private health care sectors in Hong Kong has been launched. Though potential benefits of EMR are evident, it has been a challenge to gain buy-in for the technology from private doctors (Hong Kong Legislative Council Panel on Health Services, 2009) which provide more than 70% of outpatient consultations for the population. Many of them continue to use paper-based medical records; for the few that use electronic records, most do not have data sharing capabilities (Hong Kong Legislative Council Panel on Health Services, 2009). Non-adoption of EMR is a significant concern for patients and health care organizations because its adoption is a prerequisite for benefits of the technology to be realized.

Extant literature has indicated that slow adoption or abandonment of a health information technology (IT) is likely to occur when the technology impedes routine clinical practices or when the design of the technology is not compatible with the users, tasks, or external environment (i.e., misfit between the technology and its sociotechnical system elements) (Lorenzi, Riley, Blyth, Southon, & Dixon, 1997; C. K. L. Or et al., 2009; Tan & Or, 2015, in press). For EMR systems, in particular, their implementations have failed in a number of instances because of misfit between the technology design, work practices, workflows, and clinicians' needs (e.g., Ash & Bates, 2005; Chin, 2006; DesRoches et al., 2008; Sicotte, Denis, & Lehoux, 1998). According to the sociotechnical theories and models for (health) technology implementation (Hendrick & Kleiner, 2000; C. Or, Dohan, & Tan, 2014; Sittig & Singh, 2010), to address the issue of misfit one of the necessary initial steps is pre-implementation workflow and work system assessment since it facilitates understanding of the characteristics and needs of the end-users and work units (McGowan, Cusack, & Poon, 2008; C. Or, Wong, Tong, & Sek, 2014; Samuels, 2008). By understanding the characteristics and needs, practitioners can design the technology in a way that can enhance the technology-user-workflow integration. Indeed, studies assessing workflow and work systems largely via observations for health IT implementation and adoption exist. For instance, Dykes et al. (2005) conducted an observational analysis of workflow to determine how primary care providers interacted with an EMR system. Their analysis led them to establish new adoption and training strategies that resulted in improvement in the adoption of some of the system features. Or et al. (2014) performed field observations and interviews to understand why a clinical information system of a nursing home was abandoned and how work tasks were performed in order to identify factors affecting the implementation of the system. Unertl et al. (2009) modeled the workflow, information flow, and work practices of ambulatory clinics providing chronic disease care by using observational and interview data. Their study led to the development of guidelines for health IT design to support chronic disease care.

The present study aimed to scrutinize pre-implementation clinical workflows in private primary health care clinics in order to address the following questions: 1) what were the major clinical activities in the private primary health care clinics in Hong Kong and the associated gaps and weaknesses as well as informatics challenges in the clinical process; and 2) what would be the focuses when developing an effective EMR system for the gaps, weaknesses, and challenges.

METHODS

Grounded in the theory of sociotechnical analysis of work system (Hendrick & Kleiner, 2000), this study used a direct field observation method to analyze the workflows and possible improvement opportunities.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The method was adopted because it allowed us to understand how work activities and issues occurred in real situations (Hendrick & Kleiner, 2000). As mentioned earlier, previous studies have demonstrated the use of observations to examine workflows and work systems. Based on the Systems Engineering Initiative for Patient Safety (SEIPS) model of work system and patient safety (Carayon et al., 2006), the observations focused on interactions among the individuals (i.e., health care professionals, patients and other caregivers), clinical tasks, technology, work environment, and organization. A health informatician with nursing background (KW) and a human factors researcher (IC; see the acknowledgments) were trained and took part together to observe a total of forty-three health care professionals (16 primary care physicians and 27 clinical assistants (CAs)) of fifteen clinics. The two researchers stood in a non-interfering area, such that doctor-patient encounters were not disturbed. After the observations, the researchers asked the health care professionals follow-up open-ended questions for seeking clarifications about work procedures and rationale behind actions. Those fifteen clinics were selected because they were a private primary care clinic, had not had any EMR sharing system in place or only had basic computer applications (e.g., Microsoft Excel) for data management, and were readily available for participation. Based on prior studies (Flick, 2006; C. K. L. Or et al., 2005; Poon et al., 2006; Samuels, 2008; Unertl et al., 2009), observational data of this present study was translated into a workflow diagram whereby the clinical workflow challenges were analyzed and elucidated.

RESULTS

Table 1 presents the characteristics of the study sample. Figure 1 illustrates step-by-step details of the care delivery processes, starting with patient registration and continuing through patient discharge.

In summary, CAs first check in patients upon their arrival. For a new patient's registration, the CA requires the patient to write down his/her demographic information on a piece of paper and to present his/her Hong Kong identity (HKID) card or passport as proof of identity. The CA also collects the patient's health insurance information and then verifies his/her insurance eligibility status and coverage

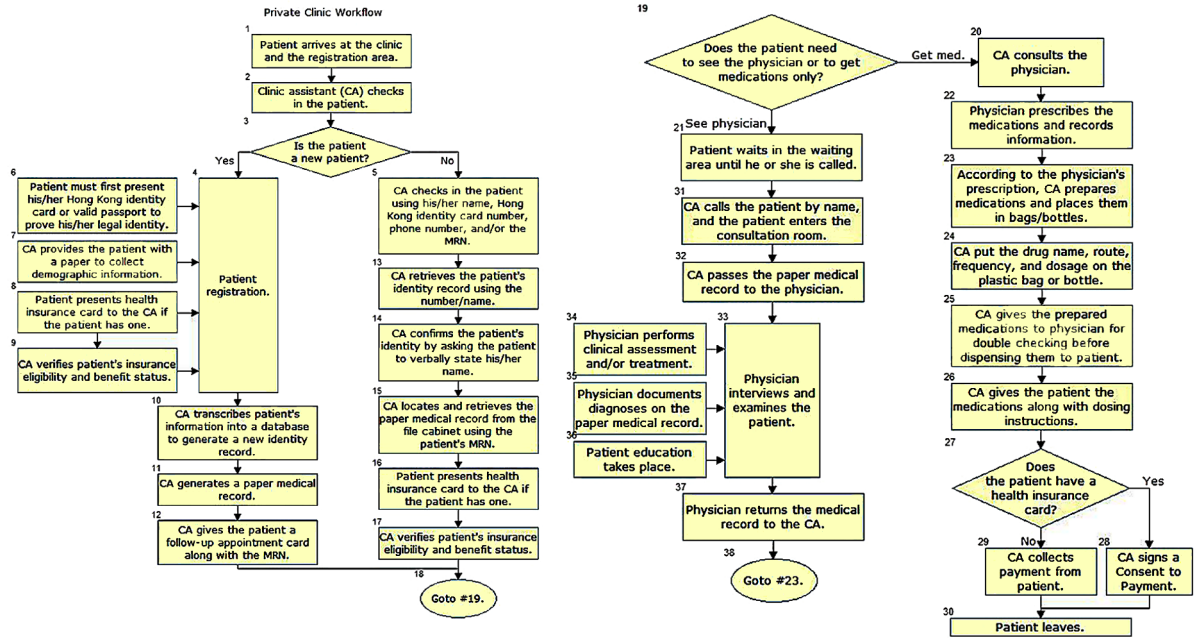
Table 1. Study sample characteristics

	Physicians (N=16)	**Clinical Assistants (N=27)**
Males/females	14/2	0/27
Mean age (SD)	47.5 (11.2) years[†]	34.1 (11.8) years
Experience in medicine Mean (SD) Range	 21.8 (11.3) years[†] 5-40 years[†]	 10.7 (9.8) years 0-40 years
Self-reported experience level in using computer None Beginner Competent Expert	 0 7 (43.8%) 9 (56.2%) 0	 0 11 (40.7%) 16 (59.3%) 0
Had experience using health IT (e.g., barcoding and standalone patient record systems)	10 (62.5%)	7 (25.9%)

[†]N=15 as one physician refused to provide the information

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Primary care workflow processes observed in private care clinics



type by calling the insurance company, using an insurance verification card reader, or through the company website. Following that, the CA manually transcribes the demographic and insurance information into an identity record that is maintained in either a paper notebook or an electronic database (e.g., an Excel spreadsheet). The CA also generates a paper-based medical record and a medical record number (MRN) for the patient (a MRN is a unique, clinic-specific number allowing identity records to be linked to medical records). For patients who visited the clinic before (old cases), they present their MRN to the CAs for registration. With a MRN, CAs can directly retrieve patients' paper-based medical records. If patients fail to provide their MRN, CAs must first search and retrieve the patients' identity records in order to obtain their MRNs before being able to locate the medical records.

Next, the CA put the patient into the consultation room for medical consultation. Meanwhile, the CA passes the patient's medical record to the physician and collects the record of last patient for preparing medications. As was observed, the physician's work involves generation and examination of different pieces of information (e.g., medical records, sick leave certificates, attendance certificates, receipts, health insurance reimbursement claim forms, and/or other medical records and diagnostic test reports brought by patients from various hospitals or health centers). Tools and technologies such as glucometers, stethoscopes, otoscopes and oximeters are used and discussions are held with patients and their family members. For each patient, the physician documents the patient's diagnosis, progress, treatment, vitals and prescriptions in a paper chart.

Sometimes patients only request to refill medications. In this case, CAs pass the patients' medical records to physicians, who review the patients' diagnosis, progress, and prescription history, and writes the prescription. CAs dispense medications accordingly and place them in bags or bottles with labels containing the patient's full name, drug name and route, dosage and frequency of intake, and safety instructions. This labeling can be completed by computer printing or by hand. Following that, the physician

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

cross-checks the medications and drug labels before the CA gives the medicines to the patient. When discharging patients, CAs collect payment from patients and for those having health insurance they also need to sign an insurance reimbursement claim form.

DISCUSSION

To the author's knowledge, this is the first study exploring in detail the clinical workflow process and informatics challenges in the private health care sector in Hong Kong—a less studied population. Many previous health IT implementation initiatives did not examine workflows and work systems prior to the implementation partly perhaps because the practitioners might believe that the technologies would automatically fit the existing work system and thus be accepted and used by users when they were implemented. This study identifies and highlights the major clinical workflow issues through the effective use of workflow and work system analysis and provides new practical insights into critical design aspects that should be addressed in the development of the EMR sharing system. More specifically, the workflow model of this present study is useful for health care organizations to evaluate design flaws in a primary care work system as well as for health informatics practitioners to develop an effective EMR system for the system design problems. As can be seen from the foregoing, there are gaps, design weaknesses and challenges in the current work system, workflow and information flow.

Overall, the process of collecting and managing patients' data is cumbersome and disorganized, entails considerably inefficient work tasks, and is vulnerable to effective information transfer and communication. First, in new patient registration, the use of scraps of paper is a frequently observed method to collect patients' information. To maintain the information, the CA creates an identity record in either an electronic or paper format by manually copying the patient's information on papers into a computer database or paper notebook, respectively. Clearly, the use of scraps of paper is an unreliable way to secure patients' data. Also, the process of manual transcription of data from one form to another can be a source of error.

Furthermore, a paper medical record and an associated MRN are also created for each patient to maintain clinical data. During old case patient check-in, the CA must first obtain the MRN from the patient and use the number to locate the patient's identity record for verifying their identity and the paper medical record for physician charting. However, as was observed, not all patients were able to recall or had their MRNs with them at the time of check-in. As a result, the CA was unable to instantly locate any of the records. Instead, the CA used other information of the patient (e.g., name, phone number, and/or HKID/passport number) to search for the identity record for the MRN. It is especially troublesome for clinics using paper-based identity records because CAs need to manually search through the paper notebook to obtain patients' MRNs. The search process is time-consuming as CAs need to flip and scan through pages of records, whereas search through an electronic database is less tedious and quicker. All the aforementioned information gathering, generation and retrieval mechanisms that consist of several inefficient and redundant steps make the care service delivery process disorganized, time-consuming and error-prone.

It is also of great concern for the clinicians to handle many different types and pieces of information from various sources and to maintain various records stored in different places in the clinics. At another instance, it was observed that the CA was unable to find the paper medical record in the file cabinet for a patient as the record was misfiled after the previous visit. The record was found only after the CA went

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

through a long series of checking and searching. Using multiple records to maintain data is unreliable and is likely prone to errors. To avoid multiple and duplicate data entry or filing, it is suggested that an EMR system be designed to enable the clinicians to acquire and manipulate all patient information (including demographic, insurance, medical, lab test and others) in one record system and to access the records by using a single, unique identification number, such as an identity card or passport number.

Information transfers between physicians, CAs, and patients are primarily by paper records and notes written on pieces of paper as it is handy. Moreover, CAs are required to walk back and forth between their work area and the consultation room when transferring documents and notes. It is suggested that EMRs can eliminate this cumbersome and time consuming step by offering an easy-to-use and fast-response mechanism that allows an efficient information transfer and communication among the physicians, CAs, and patients.

The handling of patients' health insurance information and payment reimbursement documents also makes the workflow inefficient. In case of patients having an insurance plan, CAs needs to use different tools to identify and verify patients' insurance eligibility and coverage. It is suggested that an EMR system that enables full access to updated information on patients' insurance coverage and/or that performs automatic eligibility checks should simplify the verification process.

To promote EMR adoption and implementation, a critical key is that system features and functions should match the users' needs (Karsh, Holden, & Or, 2011). To achieve this, system developers must identify the required functionality within the EMR. The present study observed that physicians were engaged in various work tasks during patient encounters. Based on the observations, an EMR system should at least have the ability for the physicians to extract patients' data from the system, such as medical history and laboratory test results; document patient diagnoses, progress, and results; record treatments provided; order laboratory tests; prescribe medications; access clinical and teaching materials; compose sickness certificates and substantiation of attendance and services; prepare billing documents, such as reimbursement claim forms and receipts; and provide electronic signature on notes and documents. The observations also suggested that the EMR system should allow the CAs to easily print medication labels directly off the machine and attach them to the drug bags and bottles as this can help reduce clerical and labeling errors caused by handwriting. On the other hand, without this information about what the physicians and CAs need in their work, it would be unlikely to develop an EMR system that would just fit the needs; as a result, the use of such system is likely to be ineffective and inefficient, and the system can eventually be abandoned.

This study has certain limitations that need to be addressed in future research. The generalizability of this study should be considered in light of the fact that it was conducted on a sample of physicians and CAs of only fifteen private primary care clinics. Further studies with a larger sample size covering clinics with different specializations are needed as workflows of specialist care could be different. The author also acknowledges that the study sample is limited to a single region. However, it has been suggested that findings from studies of this type are most relevant to jurisdictions with similar technical capabilities, geographical characteristics, historical background, socioeconomic development, or cultural identity to Hong Kong, such as Shanghai, India, Singapore and other (former) British colonies with similar health care systems (Leung, Yu, Wong, Johnston, & Tin, 2003). There were some methodological challenges we encountered in the study that might have limited the amount of data we could collect. For instance, some clinics that were invited to participate in the study were not observed because they were not available. Also, in few cases, during the data collection, the researchers were asked to leave

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the consultation rooms as the patients did not want to be observed during the consultations. However, despite the limitations, our study provides some useful baseline knowledge and significant insights into an EMR design and implementation.

CONCLUSION

This present pre-implementation study scrutinized and identified everyday clinical activities, workflows, information flow, and the stakeholders' needs in private primary health care clinics that enabled us to recognize the inadequacy of the workflow processes and to acquire a thorough insight into what to focus on when developing and implementing the EMR system. Using the knowledge about the workflows and needs can help ensure the EMR system to include effective functions and features and consequently increase the harmony among the system design, the care processes, and the users' characteristics. This harmony is an essential step to promote EMR acceptance and adoption. In general, this study leads to the following important implications for health IT implementation. Poor fit between health IT design, workflows, and work practices can lead to unintended adverse consequences, such as work practices interruptions, additional workload, and technology rejection. To improve fit pre-implementation assessments of existing workflows and user work patterns and behaviors are necessary.

Clinical Relevance Statement

Workflow and work system assessment prior to health IT implementation is critical as the workflow knowledge can facilitate understanding end-user and work system needs and thus can help increase the harmony between the technology design and the care processes.

Conflict of Interest Statement

The author declares that they have no conflicts of interest.

Protection of Human Subjects

Approval of the institutional review board of the University of Hong Kong was received, and written informed consent was obtained from participants.

ACKNOWLEDGMENT

This study was supported in part by the Hospital Authority of Hong Kong/the University of Hong Kong, contract research project RCGAS-NA100001 C. Or (PI) and the Hong Kong Research Grants Council, Theme-based Research Scheme (Grant # T32-102/14-N; PI: F. Chen). The author is grateful to the physicians and clinical assistants for their participation and thank Katie Wong, RN; Irene Cheung; Ellen Tong, RN; and Antonio Sek, MD for their assistance on data collection and experimental design.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Ash, J. S., & Bates, D. W. (2005). Factors and forces affecting EHR system adoption: Report of a 2004 ACMI discussion. *Journal of the American Medical Informatics Association*, *12*(1), 8–12. doi:10.1197/jamia.M1684 PMID:15492027

Carayon, P., Hundt, A. S., Karsh, B. T., Gurses, A. P., Alvarado, C. J., Smith, M., & Brennan, P. F. (2006). Work system design for patient safety: The SEIPS model. *Quality & Safety in Health Care*, *25*(Suppl I), i50–i58. doi:10.1136/qshc.2005.015842 PMID:17142610

Chin, T. (2006). Avoiding EMR meltdown: How to get your money's worth. *American Medical News.* Retrieved from http://www.amednews.com/article/20061211/business/312119994/4/

DesRoches, C. M., Campbell, E. G., Rao, S. R., Donelan, K., Ferris, T. G., Jha, A., ... Blumenthal, D. (2008). Electronic health records in ambulatory care–A national survey of physicians. *The New England Journal of Medicine*, *359*(1), 50–60. doi:10.1056/NEJMsa0802005 PMID:18565855

Dykes, P. C., McGibbon, M., Judge, D., Li, Q., & Poon, E. G. (2005). Workflow analysis in primary care: implications for EHR adoption. *Paper presented at the Annual Symposium of American Medical Informatics Association*.

Flick, U. (2006). *Observation and ethnography: An Introduction to Qualitative Research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Hendrick, H. W., & Kleiner, B. M. (2000). *Macroergonomics: An Introduction to Work System Design*. Santa Monica, CA: Human Factors and Ergonomics Society.

Hong Kong Legislative Council Panel on Health Services. (2009). Development of a territory-wide electronic health record sharing system. *LC Paper No. CB(2) 1006/08-09(03)*.

Karsh, B., Holden, R. J., & Or, C. K. L. (2011). Human factors and ergonomics of health information technology implementation. In P. Carayon (Ed.), *Handbook of Human Factors and Ergonomics in Health Care and Patient Safety* (2nd ed., pp. 249–264). CRC Press. doi:10.1201/b11219-23

Leung, G. M., Yu, P. L. H., Wong, I. O. L., Johnston, J. M., & Tin, K. Y. K. (2003). Incentives and barriers that influence clinical computerization in Hong Kong: A population-based physician survey. *Journal of the American Medical Informatics Association*, *10*(2), 201–212. doi:10.1197/jamia.M1202 PMID:12595409

Lorenzi, N. M., Riley, R. T., Blyth, A. J. C., Southon, G., & Dixon, B. J. (1997). Antecedents of the people and organizational aspects of medical informatics: Review of the literature. *Journal of the American Medical Informatics Association*, *4*(2), 79–93. doi:10.1136/jamia.1997.0040079 PMID:9067874

McGowan, J. J., Cusack, C. M., & Poon, E. G. (2008). Formative evaluation: A critical component in EHR implementation. *Journal of the American Medical Informatics Association*, *15*(3), 297–301. doi:10.1197/jamia.M2584 PMID:18308984

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Or, C., Dohan, M., & Tan, J. (2014). Understanding critical barriers to implementing a clinical information system in a nursing home through the lens of a socio-technical perspective. *Journal of Medical Systems*, *38*(9), 1–10. doi:10.100710916-014-0099-9 PMID:24395031

Or, C., Wong, K., Tong, E., & Sek, A. (2014). Private primary care physicians' perspectives on factors affecting the adoption of electronic medical records: A qualitative pre-implementation study. *Work*, *48*(4), 529–538. PMID:24346272

Or, C. K. L., Casper, G., Karsh, B., Brennan, P. F., Burke, L. J., Carayon, P., . . . Sebern, M. (2005). Work system analysis of home nursing care and implications for medication errors. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1052-1056).

Or, C. K. L., Valdez, R., Casper, G., Carayon, P., Burke, L. J., Brennan, P. F., & Karsh, B. (2009). Human factors and ergonomics in home care: Current concerns and future considerations for health information technology. *Work*, *33*(2), 201–209. PMID:19713630

Poon, E. G., Blumenfeld, B., Hamann, C., Turchin, A., Graydon-Baker, E., McCarthy, P. C., ... Broverman, C. A. (2006). Design and implementation of an application and associated services to support interdisciplinary medication reconciliation efforts at an integrated healthcare delivery network. *Journal of the American Medical Informatics Association*, *13*(6), 581–592. doi:10.1197/jamia.M2142 PMID:17114640

Samuels, C. (2008). Electronic health records implementation: workflow. In J. H. Carter (Ed.), *Electronic Health Records: A Guide for Clinicians and Administrators* (2nd ed., pp. 453–472). Washington, DC: American College of Physicians.

Sicotte, C., Denis, J. L., & Lehoux, P. (1998). The computer based patient record: A strategic issue in process innovation. *Journal of Medical Systems*, *22*(6), 431–443. doi:10.1023/A:1020674527409 PMID:9871877

Sittig, D. F., & Singh, H. (2010). A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Quality & Safety in Health Care*, *19*(Suppl. 3), i68–i74. doi:10.1136/qshc.2010.042085 PMID:20959322

Tan, J., & Or, C. (2015). Themed issue on matching policies and socio-technical challenges to enhance the uptake of health information technologies. *Health Policy and Technology*, 4(4), 297–298. doi:10.1016/j.hlpt.2015.09.002

Unertl, K. M., Weinger, M. B., Johnson, K. B., & Lorenzi, N. M. (2009). Describing and modeling workflow and information flow in chronic disease care. *Journal of the American Medical Informatics Association*, *16*(6), 826–836. doi:10.1197/jamia.M3000 PMID:19717802

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 10(4); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 56-64; copyright year 2015 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 65
Extraction of Protein Sequence Motif Information using Bio-Inspired Computing

Gowri Rajasekaran
Periyar University, India

Rathipriya R
Periyar University, India

ABSTRACT

Nowadays there are many people affected by the genetic disorder, hereditary diseases, etc. The protein complexes and their functions are detected, in order to find the irregularity in the gene expression. In a group of related proteins, there exist some conserved sequence patterns (motifs) either functionally or structurally similar. The main objective of this work is to find the motif information from the given protein sequence dataset. The functionalities of the proteins are ideally found from their motif information. Clustering approach is a main data mining technique. Besides the clustering approach, the biclustering is also used in many Bioinformatics related research works. The PSO K-Means clustering and biclustering approach is proposed in this work to extract the motif information. The Motif is extracted based on the structure homogeneity of the protein sequence. In this work, the clusters and biclusters are compared based on homogeneity and motif information extracted. This study shows that biclustering approach yields better result than the clustering approach.

INTRODUCTION

Protein Sequence

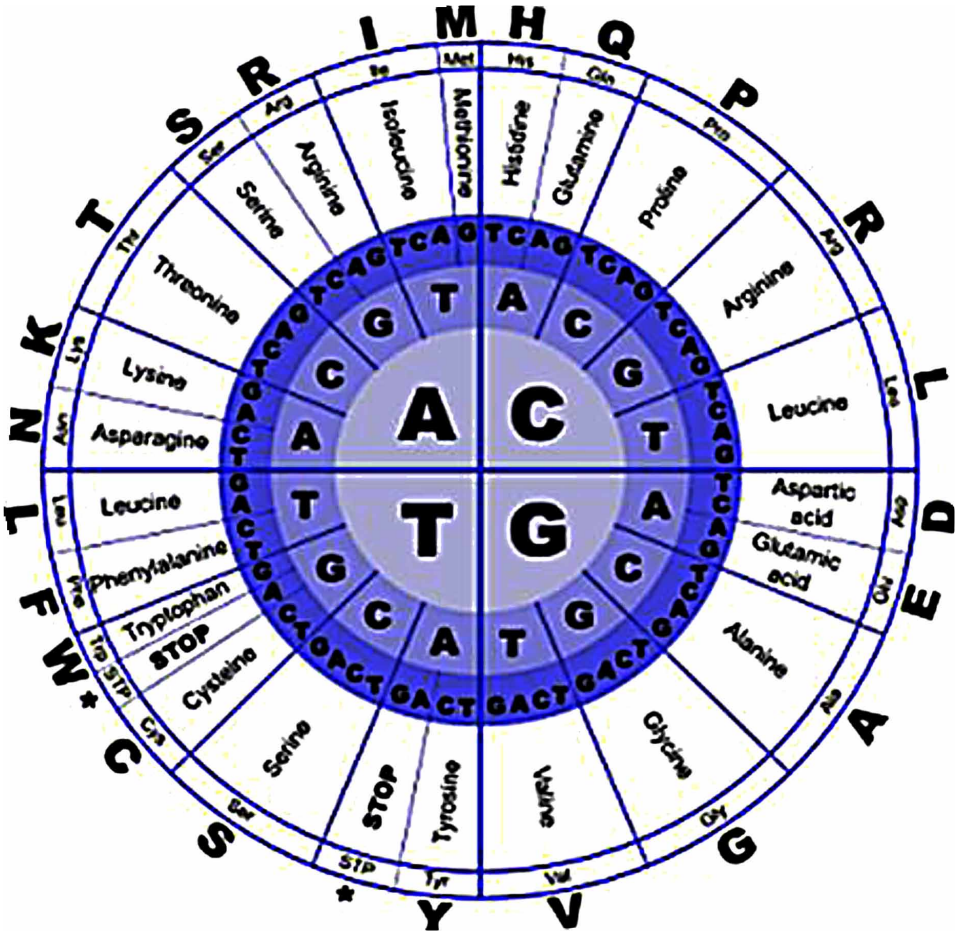
Proteins (Vincent, Bernard, & SinanKockara, n.d.) are present in every cell of the organisms. They are involved virtually almost in all cellular activities. They are responsible for the various metabolic activities, nutrition transportation, regulations and etc. They exist as single chain molecule, as a three dimensional

DOI: 10.4018/978-1-7998-1204-3.ch065

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. The IUPAC code for PROTEIN from DNA code



structures or even in the bundle or complex forms. The protein plays a vital role in cellular processes. The protein consists of twenty amino acids. They possess different characteristics such as hydrophobic, hydrophilic, polar, non-polar and etc. It is the great challenge to the bioinformatics that to find which combination of proteins are responsible for what kind of activities. The structure and function discovery of proteins in living organisms is vital role in understanding the background of various cellular processes. It is helpful in treating various diseases, in detecting the drugs to peculiar diseases.

The irregularity in the proteins can be found if their actual functionality and their protein complex are known. The biologist will interpret the functionality of the protein complex based on their chemical properties.

These significant protein complexes of various organisms are discovered using the computational techniques. Protein complexes are an assembly of proteins that build up some cellular machinery; commonly spans a dense Sub-network of proteins in a protein interaction network. The gene expression would not help much in detecting the functionalities of the gene. Protein sequence can be generated from the DNA/mRNA sequence that codes for the protein are shown in the Figure 1. The four basic amino acid combines to form the protein that is shown. The 20 different amino acids that are present in the protein sequence along with their chemical name are listed in the Table 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Amino acid codes

1-LETTER	3-LETTER	DESCRIPTION
A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartic acid
C	Cys	Cysteine
Q	Gln	Glutamine
E	Glu	Glutamic acid
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine
B	Asx	Aspartic acid or Asparagine
Z	Glx	Glutamine or Glutamic acid

As the protein sequences are in the form of long amino acid sequence, the repeated protein groups are not found manually. These protein groups are denoted as motifs. The protein structure is of many types secondary, tertiary, quaternary, 3 fold and so on. The secondary structures are used in this work. The structural similarity of the detected protein complexes is used to extract the homologous complexes. The functionality of proteins is discovered by various methods like sequence-motif based method, homology based methods, and structure based methods. The motifs can be extracted from the clusters that are generated by various computational techniques.

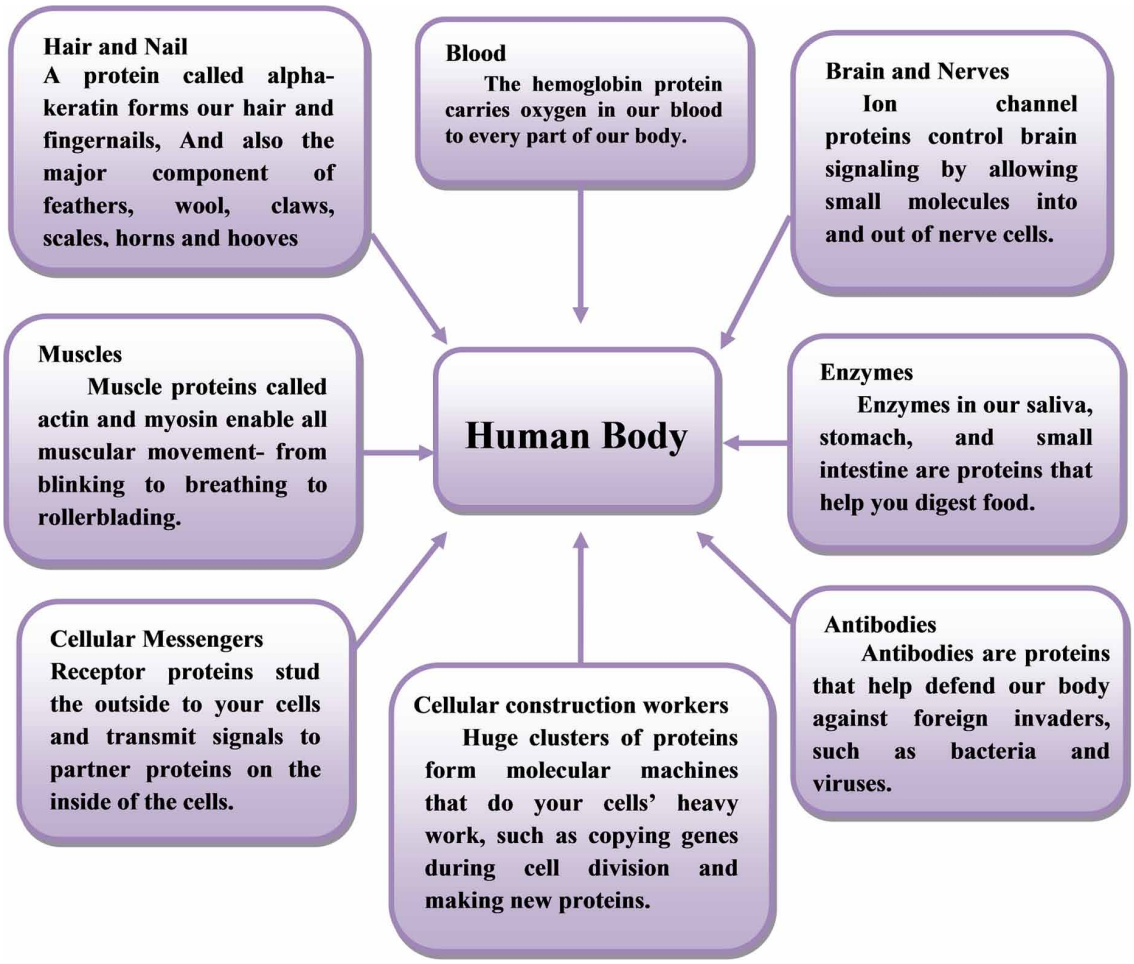
Roles of Proteins

The proteins[1] are responsible for various activities of the biological system. They present in various parts of the human body. The Figure 2 shows the basic roles of protein in the human body. Besides this, the protein also has various other roles. The motifs are helpful in predicting these functionalities of the protein. The proteins are present in

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Blood**: The hemoglobin protein carries oxygen in our blood to every part of our body.
- **Hair and Nail**: Protein called alpha-keratin forms our hair and fingernails, and also the major component of feathers, wool, claws, scales, horns and hooves
- **Muscles**: Muscle proteins called actin and myosin enable all muscular movement- from blinking to breathing to rollerblading.
- **Cellular Messengers**: Receptor proteins stud the outside of your cells and transmit signals to partner proteins on the inside of the cells.
- **Cellular Construction Workers**: Huge clusters of proteins form molecular machines that do your cells' heavy work, such as copying genes during cell division and making new proteins.
- **Brain and Nerves**: Ion channel proteins control brain signaling by allowing small molecules into and out of nerve cells.
- **Enzymes**: Enzymes in our saliva, stomach, and small intestine are proteins that help you digest food.
- **Antibodies**: Antibodies are proteins that help defend your body against foreign invaders, such as bacteria and viruses.

Figure 2. Various roles of protein in human body



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Clustering

Clustering (Berkhin, 2002) is one of the known data mining technique used to group similar kind of data elements. It is used to discover similar patterns from a sea of data. The similarity among objects in the same cluster is greater than in different clusters. It is one of the unsupervised learning algorithms. The clustering (DEIB, n.d) is helpful in intrinsic grouping the unknown data based on our objective. It is a main task of exploratory data mining (Wikipedia, n.d.), and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and Bioinformatics. The motif information is extracted using various techniques like clustering with K-Means, hybrid K-Means, Self-Organizing Maps, etc., in the literature. As all these areas are dealing with the unclassified data, the clustering is well suited to these kinds of research areas. The clusters can be found based on various similarities between the data, such as intra distance and inter distance of the clusters. The quality of clusters will be evaluated based on our objective.

There are a number of problems with clustering. Some of them are as follows.

- Current clustering techniques do not address all the requirements adequately (and concurrently).
- Dealing with large number of dimensions and large number of data items can be problematic because of time complexity.
- The effectiveness of the method depends on the definition of "distance" (for distance-based clustering).
- If an obvious distance measure doesn't exist, we must "define" it, which is not always easy, especially in multi-dimensional spaces.
- The result of the clustering algorithm (that in many cases can be arbitrary itself) can be interpreted indifferent ways.

Biclustering

Biclustering (Madeira & Oliveira, 2004) is another data mining technique. It is also named as Co-Clustering, Two Way Clustering. It generates biclusters of different sizes and characteristics. The process of grouping data based on both the rows (segments) and columns (amino acids). Biclustering is useful for Bioinformatics and many other research areas. It is NP-Complete.

The major difference (Berkhin, 2002; Madeira & Oliveira, 2004) between the clustering and biclustering are as follows.

1. Clustering applied to any one of the dimensions of the dataset, but biclustering applies to both dimensions.
2. Clustering derives the global model, biclustering derives the local model.
3. Biclustering groups more similar patterns than the clustering process.

Contribution

The contributions of the thesis are as follows.

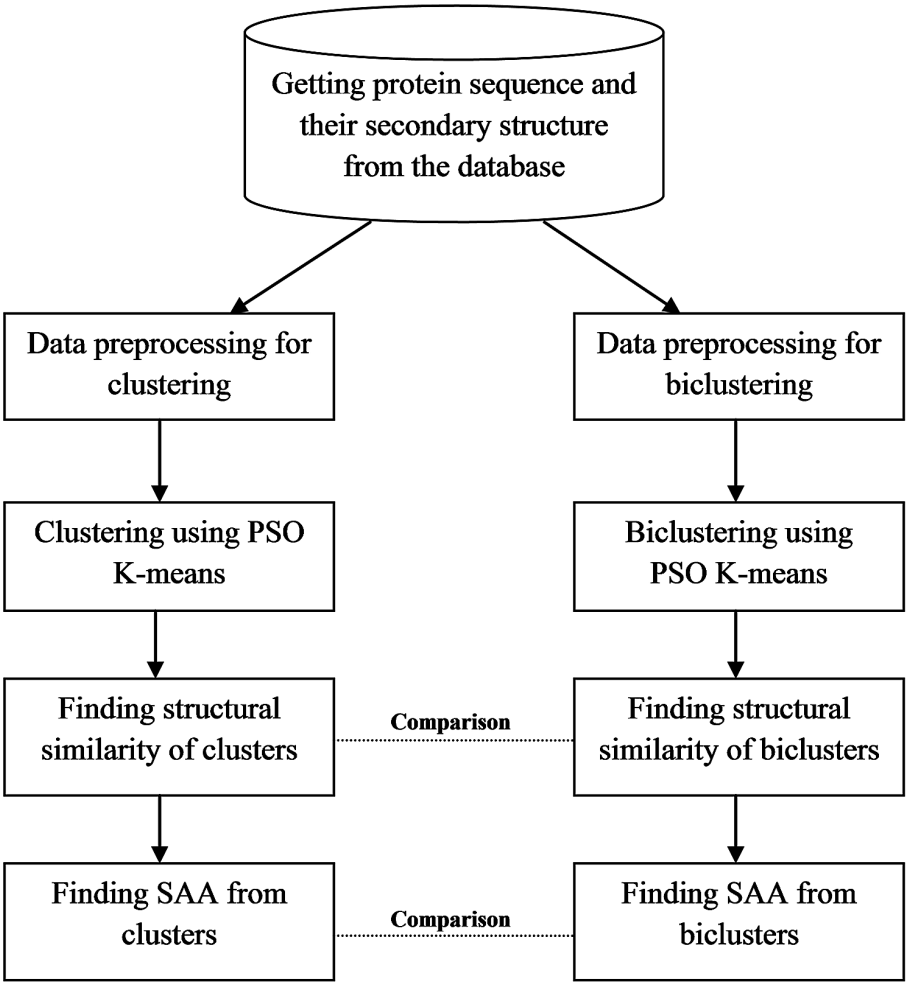
1. Studying various methods of extracting protein motifs from the literature.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. The proposed work is to extract the motif from the protein by clustering the given protein sequence with PSO K-Means.
3. The proposed work is to extract the motif from the protein by biclustering the given protein sequence with PSO K-Means.
4. Comparison between the quality of the clusters and biclusters based on structural similarity.
5. Comparison between significant amino acids extracted from clusters and biclusters.
6. The overview of the proposed work is shown in the Figure 3.

The protein sequence is extracted from the databases that are available for the protein sequence and motifs such as BLOCKS (Pietrokovski, Henikoff, & Henikoff, 1996), PROSITE (Bairoch, 8ucher & Hofmann, 1996), PISCES (Wang & Dunbrack, 2003) and so on. The protein sequence is preprocessed for both clustering and biclustering processes. The clusters and biclusters are compared based on their structural similarity, extracted significant amino acids and protein motifs extracted.

Figure 3. Overview of the proposed work



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

In this work (Zhaoa-Xing-Ming, 2005), they have proposed a novel approach for extracting features from motif content and protein composition. They have used Support Vector Machine (SVM) along with the combination of Genetic Algorithm (GA) and Principal Component Analysis (PCA) for extracting the protein motif.

In this work (Fujiwara & Konagaya, 2008), they have proposed a methodology for protein extraction in their work, they have used the Hidden Markov Model (HMM) to detect the protein motifs from the proteins with motifs unidentified. They train the HMM with test samples of protein motifs of various categories to detect the protein motifs as similar to the trained dataset.

In this work (Chen, Pellicer, Tai, Harrison, & Pan, 2009), the authors have proposed a novel efficient granular computing models for protein sequence motif and structure information discovery. The protein sequence motifs are extracted using the granular computing with the Fuzzy improved K-Means model and Fuzzy greedy K-Means model.

In this work (Cheng, Hudson, Kim, Crawford, Wright, & Che, 2011), they have defined another method for protein sequence motif extraction using Decision Forest. They have used the decision forest combined with fuzzy logic and granular computing to get more efficient results than their previous work.

In this work (Vincent, Bernard, & SinanKockara, n. d.), the authors have proposed a biclustering algorithm for extracting the protein motifs. They obtained the clusters using Fuzzy Greedy K-Means (FGK) model along with the Granular Computing. They used Cheng and Church methodology to shrink the clusters in order to get high quality biclusters.

In this work (Elayaraja, Thangavel, Ramya, & Chitralegha, 2011), the authors have devised a method for motif extraction process using the rough K-Means. They have used Rough K-Means to yield the better results than simple K-Means in their work, "Extraction of Motif Information from Protein Sequences using Rough K-Means" in 2011

In this work (Elayaraja, Thangavel, Chitralega, & Chandrasekhar, 2012), they have further suggested another method for motif extraction process using the SVD in addition to the rough K-Means in order to attain better results. Their work is "Extraction of Motif Patterns from Protein Sequences Using SVD with Rough K-Means Algorithm" in 2011

In this work (Chang, 2012), they have suggested a method for Protein Motif Extraction Using Neuro-Fuzzy optimization in 2002, they attempted to improve the speed and flexibility of protein motif identification. They proposed an algorithm to extract both rigid and flexible protein motifs and to detect consensus pattern, or motif, from a group of related protein sequences.

In this work (Chen, KripamoyAguan, Yang, Wang, Pal, & Chung, 2011), they have suggested a method for the discovery of Protein Phosphorylation Motifs through Exploratory Data Analysis, they proposed an F-Motif algorithm for motif extraction using clustering methods in 2011.

The various works in the literature show that the different hybrid methods were used for extracting the Motif Information. Those methods are any combination of Support Vector Machine (SVM), hidden Markov Model (HMM), Neural Networks, Fuzzy system, Rough set, Clustering technique and so on. Some authors were used optimization technique to improve the extraction process, such as Genetic Algorithm (GA). From the literature, it shows the combination of optimization technique can improve the results. The PSO is one of the best optimization techniques that does not stuck with local optimum. The

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

clustering technique is better suits for unidentified data. The biclustering further improves results than the clustering process, the combination of PSO with clustering and biclustering on protein sequences is proposed and yield better results. This proposed work is further discussed in the following chapters.

METHODS AND MATERIAL

The protein motifs are extracted using various data mining techniques. In this study, clustering and biclustering technique are used. The K-Means algorithm is used for this clustering process. The PSO K-Means algorithm is proposed for the protein sequence to extract motif information.

K-Means

The K-Means (Ahmadyfard & Modares, 2008) algorithm is one of the simplest, ever using and known clustering algorithms. It is suitable for many research areas. K-Means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. It is numerical, unsupervised, non-deterministic and iterative clustering method.

The given dataset is partitioned into k clusters. The data points are assigned randomly to the clusters resulting in clusters that have roughly the same number of data points. This algorithm works faster for the larger dataset. It produces the tighter clusters than hierarchical clustering. The simple K-Means algorithm is shown in the Figure 4.

Particle Swarm Optimization

Optimization is the process of determining the best possible or desirable solution. There are a number of nature inspired optimization algorithms that imitate the natural behaviors. Swarm intelligence (2006) is devised based on the collective behavior of the decentralized and self-organized system of the natural swarms. The particle swarm optimization (AshaGowdaKaregowda & SeemaKumari, 2013) is one of the swarm intelligence techniques. It is a stochastic optimization technique. It is a population based method that is designed based on the birds flocking behavior. Selecting the initial population is decided based

Figure 4. Simple K-Means algorithm

1. Select k objects from the search space represented by the objects that are to be clustered. These are initial cluster centroid.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the centroid.
4. Repeat Steps 2 and 3 until the centroid no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

on the problem criteria. The steps involved in PSO algorithm are given in the Figure 5. The number of individuals in the population is decided based on the objective of the problem. The particle at each time interval updates its position and velocity based on its own best position achieved so far and the global best position achieved by any one of the particles in its swarm. The random terms are used to weigh the acceleration of the particle. The best position is evaluated based on the objective function of the problem. The basic particle swarm optimization algorithm (Vora, 2013; Riccardo, James, & Tim, 2007) is in 0.

The velocity of the particle is updated based on equation 1, here the first term is based on the velocity of the previous iteration t, second term is cognition part, based on the local best position of the particle and the third term is social part, based on the global best position of the particle. The w represents the inertia weight of the previous velocity of the particle; it is calculated using the equation 2. The c1 and c2 are constant values chosen randomly, the R1 and R2 are random numbers where R1, R2 [0,1].

$$vel_{ij}\left(t+1\right) = wv_{ij}\left(t\right) + c_1 R_1\left(pbest_{ij} - x_{ij}\left(t\right)\right) + c_2 R_2\left(gbest_{ij} - x_{ij}\left(t\right)\right) \tag{1}$$

where,

$$w = W_{max} - \frac{W_{max} - W_{min}}{iter\left(max\right)} - iter\left(x\right) \tag{2}$$

The w_{max} is the initial weight of the particle and w_{min} is the final weight of the particle. The position of the particle at each iteration is updated using the equation 3 based on the updated velocity of the particle.

$$x_{ij}\left(t+1\right) = x_{ij}\left(t\right) + v_{ij}\left(t+1\right) \tag{3}$$

Homology Secondary Structure of Proteins

HSSP (Homology-derived Secondary Structures of Proteins) (Reinhard, Antoine & Chris, 1996) is a derived database merging information from three- dimensional structures and one- dimensional sequences of proteins.

HSSP file of any protein sequence will contain the following parts:

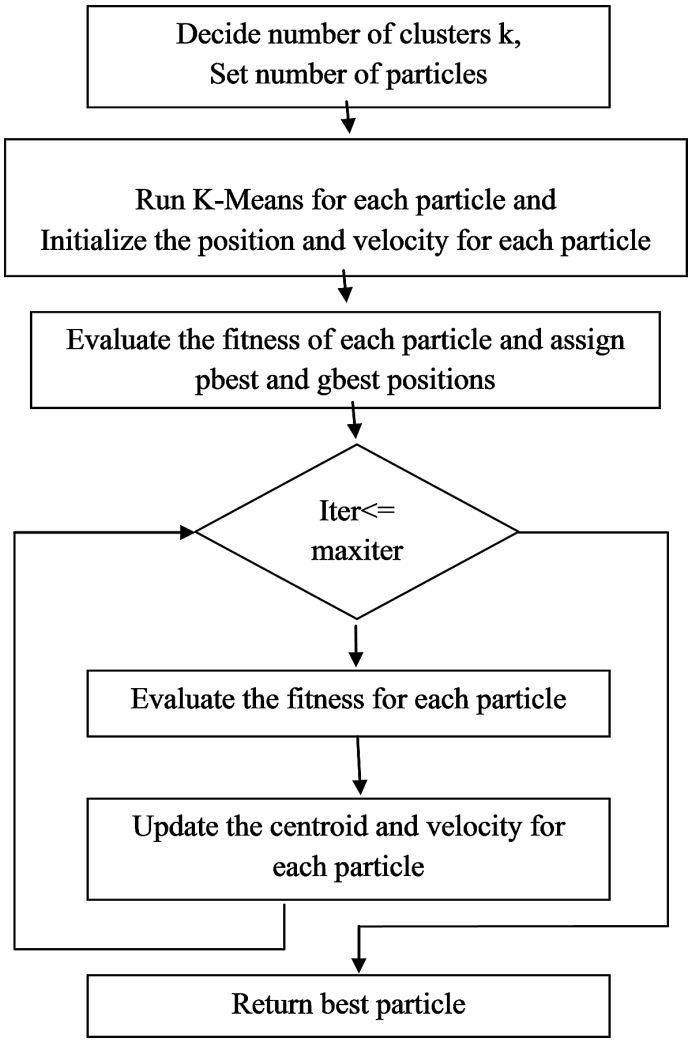
- Header information
- Protein sequence
- Alignment information
- Sequence profiles

Distance Measure

The city block distance (Vincent, Bernard, & SinanKockara, n. d.; Elayaraja, Thangavel, Chitralega & Chandrasekhar, 2012) is suitable for this type of data. Since this measure deals with n-dimensional data element, it is suitable for this type of sequence profiles. The distance measure is calculated using the equation 4.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Particle swarm optimization algorithm



$$\text{distance} = \sum_{i=1}^{R}\sum_{j=1}^{A} V_k(i,j) - V_c(i,j) \tag{4}$$

where R is the number of residues and A is the number of amino acids (20). The Vk(i,j) is the value of the profile matrix in row i and column j of segment k. The Vc(i,j) is the value at row i and column j of the centroid of the cluster c.

Cluster Validity Measure

The clusters detected by each particle are evaluated for their fitness. The intra cluster distance (Ahmadyfard & Modares, 2008) is used as the cluster validity measure. The intra distance of the cluster shows how close the elements of the cluster are placed. The average intra distance value of each particle is calculated. The particle with the minimum distance is taken as the best particle.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$\text{intra distance} = \sum_{i=1}^{n} \text{distance}(A_k\left(i\right), \text{cent}_k) \tag{5}$$

where distance(A_k(i),cent$_k$) is calculated using the city block distance as defined earlier in the equation 5.

Bicluster Validity Measure

The biclusters detected by each particle are evaluated for their fitness. The Mean Square Residue (MSR) (Madeira & Oliveira, 2004) value is used to validate the biclusters of each particle. The particle with minimal MSR value is considered as best particle. It is calculated using the equation 6.

$$\text{MSR}\left(b\right) = \frac{1}{\left|I\right|\left|J\right|} \sum_{i\epsilon I,\ j\epsilon J} \left(b_{ij} - b_{iJ} - b_{Ij} + b_{IJ}\right)^2 \tag{6}$$

Where MSR(b) is the mean square residue of a bicluster b with I rows and J columns, b_{iJ} represents the column average and b_{Ij} represents the row average.

Structural Similarity Measure

Similarity measure for the protein sequences is calculated based on the secondary structure of the protein sequence. A cluster's or bicluster's average structure (Vincent, Bernard, & SinanKockara, n.d.) is calculated using the equation 7.

$$\text{Structural similarity} = \frac{\sum_{i=1}^{ws} \max\left(P_{i,H},\ P_{i,E},\ P_{i,C}\right)}{ws} \tag{7}$$

Where ws is the window size and $P_{i,H}$ shows the frequency of occurrence of helix among the segments for the cluster in position i. $P_{i,E}$ and $P_{i,C}$ are defined in a similar way. If the structural homology for a cluster or bicluster exceeds 70%, then it can be considered as structurally identical. If the structural homology for the cluster or bicluster is between 60% and 70%, then it can be considered weakly structurally homologous.

PROPOSED METHOD: MOTIF EXTRACTION USING PSO K-MEANS

PSO K-Means

Initially the numbers of groups (k) have to be given data. This will differ for different dataset. The k value plays vital role in the K-Means, the k value that suits the given dataset can be estimated by the trial and error process. Then decide the number of particles to perform our clustering process to find the optimal result. The initial velocity and position of the particle can be set based on the objective.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Each particle will cluster the source data into k clusters. The centroid of their clusters is taken as the position of that particle. Initially, each particle will take the random position and random velocity. Then the fitness of each particle is calculated. Then for each particle the pbest position is evaluated, the velocity and position of each particle are updated until the termination criteria are met. The gbest position is updated at each iteration. The various steps involved in PSO K-Means is shown in the Figure 6.

Clustering Using PSO K-Means

The K-Means algorithm is one of the simplest, ever using and known clustering algorithms. K-means algorithm highly depends on the initial state and converges to the local optimum solution. So the optimization method is used to obtain the optimum result of the K-Means algorithm. PSO (Kunik, Solan, Edelman, Ruppin & Horn, 2005) is ideal for identifying the global optimal cluster i.e. global significant amino acid. This method suffers from local minima problem which can be eliminated by the optimization technique like PSO (Youcheng et al., 2012; AshaGowdaKaregowda & SeemaKumari, 2013)

The number of particles is set as per the user's need not greater than 100. It is set in order to overcome the complexity of the computation. The initial velocities of each particle are set to random values. The position of each particle is the cluster centroid. The number of cluster taken is fixed for each particle. The cluster intra distance is taken as the fitness in the PSO algorithm ($Fitness_{clus}$) using equation 8.

$$Fitness_{clus} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{n} P_{ik} - P_{ibest}}{K} \tag{8}$$

The particle with minimum fitness value is taken as the best particle. In the equation 4.1K represents the number of clusters, P_{ik} is the position of the i^{th}particle in the cluster k, P_{ibest} is the personal best of the i^{th}particle and n is the number of particles. The various steps involved in PSO K-Means are shown in theFigure 6.

The pseudo code for psokmeansprocedureis given in theFigure 7. the preprocessed sequence (A), number of clusters required (k) and number of particles (n) to generate are given as input to the this procedure. The procedure kmeans(A, k) is used to cluster the given dataset A into k clusters for each particle with different centroid. The max_iter is the maximum number of iterations to be performed. The evaluateFitness(pop_i) is used to evaluate the fitness value of the i^{th} particle using the equation 4.1. The evaluatePBest(p_{besti}) is used to find the local best (pbest) position of the i^{th} particle. The updateVelocity(vel_i)

Figure 6. Steps in PSO K-Means

1. Generate the initial population by choosing number of individuals in the population.
2. For each particle initialize the parameters like velocity and position.
3. While the iteration <= iteration max or termination condition
    a. Evaluate the fitness of each particle position and Update the 'pbest' and 'gbest' positions.
    b. Update the particle velocity and position using eq. 1 & 3.
4. End

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Pseudocode for PSO K-Means procedure

```
Procedure psokmeans
Input :
        A   //preprocessed dataset
        n   // number of particles
        k   // number of clusters
Output :
        gbest   //global best particle
start procedure
        for each particle i
                // call kmeans for each particle
            pop_i=kmeans(A,k)
            pbest_i=evaluateFitness(pop_i)
            vel_i=rand val
        end for
        gbest=pbest_1
            //initialize gbest with pbest of first particle
        while iter<=max_iter
            for each particle i
                evaluateFitness(pop_i)
                pbest_i = evaluatePBest(pop_i)
                vel_i = updateVelocity(pop_i)
                updatePosition(pop_i)
            end
            gbest=evaluateGBest(pbest)
        end while
        return gbest
end procedure
```

procedure is used to update the velocity of the i^{th}particlein the current iteration. The updatePosition(pop_i) procedure is used to update the position of the i^{th}particle in the current iteration. The global best (gbest) value is evaluated at the end of each iteration using evaluateGBest(pbest) procedure.

Biclustering Using PSO K-Means

Biclustering is viewed as an optimization problem with the objective of finding the biclusters with minimum MSR and high volume. PSO is ideal for identifying the global optimal bicluster i.e. global optimal sequence motif. This method suffers from local minima problem which can be eliminated by the optimization technique like PSO. A Biclustering (Das & Idicula, 2010; Rathipriya, Thangavel & Bagyamani, 2011; Cui & Potok, 2005) algorithm along with Particle Swarm Optimization (PSO) technique for protein sequence is proposed. The main objective of this algorithm is to identify the minimal subset of amino acids with coherent in their presence in the sequences. These biclusters have low MSR value, where MSR value is the quality measure of correlation between various residues and amino acids. PSO is initialized with initial biclusters which is obtained by using PSO K-Means on both the dimensions of the normalized protein frequency matrix A. This result in faster convergence compared to random particle initialization of the PSO and it also maintains high diversity in the population. Each particle of PSO (Rathipriya, Thangavel & Bagyamani, 2011; Thangavel, Bagyamani & Rathipriya, 2011) explores a possible solution. It adjusts its flight, according to its own and its companions flying experience.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The personal best (pbest) position is the best solution found by the particle during the course of the flight. This is denoted by the pbest and the optimal solution is attained by the global best (gbest). BPSO updates iteratively the velocity of each particle towards the pbest and gbest. For finding an optimal or near optimal solution to the problem, PSO keeps updating the current generation of particles. This process is repeated until the stopping criterion is met or maximum number of iterations is reached. The MSR is taken as the fitness of the bicluster ($Fitness_{bic}$) using the equation (9).

$$\text{Fitness}_{\text{bic}} = \frac{1}{|I||J|} \sum_{i \epsilon I, j \epsilon J} \left(b_{ij} - b_{iJ} - b_{Ij} + b_{IJ} \right)^2 \tag{9}$$

Where fitness is the mean square residue of a biclusterb with I rows and J columns, b_{iJ} represents the column average and b_{Ij} represents the row average.

The pseudo code for psokmeansbic procedure is given in the Figure 8 9. the preprocessed sequence (B), number of row clusters (rk) and column clusters (ck) required and number of particles (n) to generate are given as input to this procedure. The procedure kmeansbic(B, rk, ck) is used to cluster the rows of given dataset B into rk clusters and cluster the columns of given dataset B into ck clusters, then combines the row clusters and column clusters into sub-clusters known as biclusters for each particle with different centroid.

The max_iter is the maximum number of iterations to be performed. The evaluateFitness($popbic_i$) is used to evaluate the fitness value of the i^{th} particle. The evaluatePBest($pbestbic_i$) is used to find the local best (pbestbic) position of the i^{th} particle. The updateVelocity(vel_i) procedure is used to update the velocity of the i^{th} particle in the current iteration. The updatePosition($popbic_i$) procedure is used to update the position of the i^{th} particle in the current iteration. The global best (gbestbic) value is evaluated at the end of each iteration using evaluateGBest(pbestbic) procedure.

Experimental Analysis

The proposed work is implemented in MATLAB 2012, the system has 3 GHz Intel core processor, 2GB RAM and Windows7 Operating System. This section discusses the experimental analysis of each step involved in this proposed work.

Protein Sequence Dataset

In this empirical study, more than 2710 protein sequences are taken from the Protein Sequence Culling Server (PISCES) (Wang & Dunbrack, 2003) as the dataset. In this database, none of the sequences shares more than 25% sequence identity.

The frequency profiles from the HSSP is constructed based on the alignment of each protein sequence from the protein data bank (PDB) where 300 sequences are considered as homologous in the sequence database. The Sliding windows with nine consecutive residues are taken. Each window contains one sequence segment of nine continuous positions. In addition, all the sequences are considered homologous in the sequence database. Secondary structures are also taken from DSSP. DSSP is a database of secondary structure assignments for all protein entries in the PDB. The various parameters used in our work are listed in the Table 2. The 30 sample protein sequences are taken and it is converted into 10,635 segments.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Pseudocode for psokmeansbic

```
Procedure psokmeansbic
Input :
        B   //preprocessed dataset for biclustering
        n    // number of particles
        rk    // number of row clusters
        ck     //number of column clusters
Output :
        gbestbic    //global best particle

start procedure
        for each particle i
                popbic_i=kmeansbic(B, rk, ck)
                        // call kmeansbic for each particle
                pbestbic_i=evaluateFitness(popbic_i)
                vel_i=rand val
        end for
        gbestbic=pbestbic_1
                //initialize gbest with pbest of first particle
        while iter<=max_iter
                for each particle i
                        evaluateFitness(popbic_i)
                        pbestbic_i = evaluatePBest(popbic_i)
                        vel_i = updateVelocity(popbic_i)
                        updatePosition(popbic_i)
                end
                gbestbic=evaluateGBest(pbestbic)
        end while
        return gbestbic
end procedure
```

Table 2. List of parameters

Number of sequences	30
Number of segments	10,635
Number of residues	9
Number of clusters	50-100
Number of biclusters	Row clusters: 50-100 Column clusters: 3-5
Number of particles	30-100

Sequence Representation

Each sequence is represented using sliding window (Vincent, Bernard & SinanKockara, n.d.; Elayaraja, Thangavel, Chitralega & Chandrasekhar, 2012). A sliding window contains residues as rows and amino acids as columns. In this study nine consecutive residues are taken. Normally, twenty Amino Acids are used to represent the protein sequence. The size of each sequence differs from the other. In order to make

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

all the sequence of same size and to convert the categorical data into numerical the following steps are performed. First arrange the one-dimensional sequence data into two-dimensional by taking the HSSP profile frequency values of every nine consecutive residues as one segment. Now each segment will be in the size nine by twenty. This preprocessing step is shown in the Figure 9. The secondary structures of the protein sequence are collected from the DSSP database. There are eight classes in secondary structure which is mapped to three classes the following conversion model: assigning H, G, and I to H (Helices), assigning B and E to E (Sheets), and assigning all others to C (Coils).

These sliding windows are used for clustering process, whereas for biclustering each window is normalized to a single row. For this purpose, the normalization techniques like average value, difference between max and min value and most frequent value can be taken. Now the data become two-dimensional with each row representing a segment and the column represents the average composition of the 20 Amino acids in that segment. This preprocessing step is shown in the Figure 10.

Figure 9. Motif comparator preprocessing protein sequence for clustering

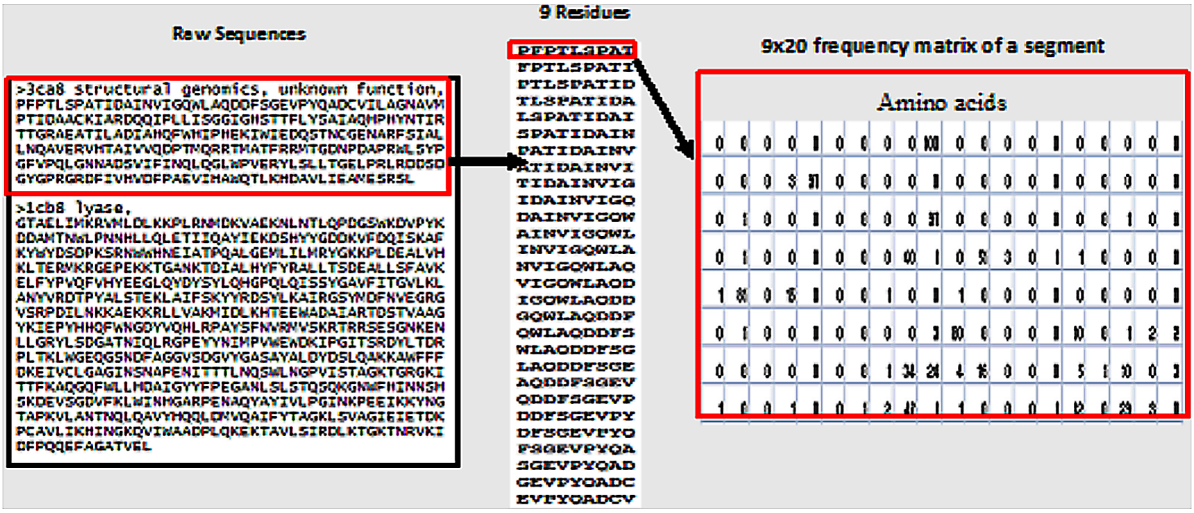
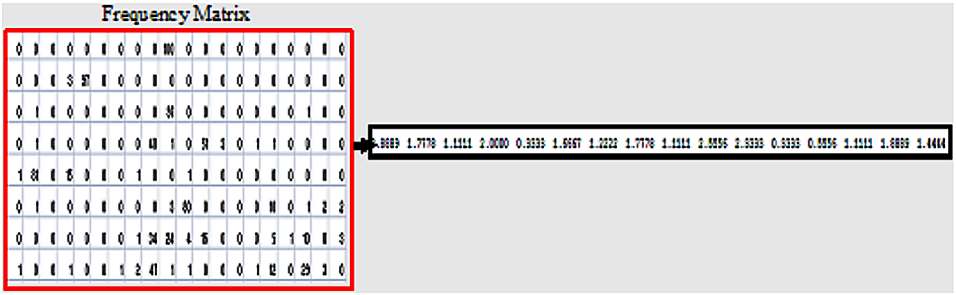


Figure 10. Motif comparator preprocessing protein sequence for biclustering



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

RESULTS AND DISCUSSIONS

The result of both the clustering and biclustering process using PSO K-Means model as proposed in this work are tabulated and interpreted below. The comparison between the clustering and biclustering process is performed based on their secondary structure similarity is discussed. The significant amino acids of the protein sequence are extracted and visualized using logo representation. The motifs detected for the given sequences are taken as superset for the significant amino acids and their interpretation is also tabulated.

Structural Similarity Comparison

The clustering technique is quite inefficient when compared to the biclustering process. This is proven from the secondary structure similarity value. The number of clusters (Vincent, Bernard, & SinanKockara, n. d.) with higher similarity value generated by our model is less when compared to the number of biclusters generated.

The Table 3 shows that biclustering can detect the biclusters with more similarity value. So the biclustering technique is better when compared to the clustering technique for protein motif extraction process. The structural similarity value is calculated using the equation 7. The clusters or biclusters with similarity value greater than 70% is taken as highly homologous and with a similarity greater than 60% and less than 70% is taken as a weakly homologous sequence. The Figure 11 is the plot with the number of clusters and biclusters with structural similarity as given in the Table 3. From this figure, it is clear that the biclustering model can detect more homologous sequence than the clustering model.

Sequence Motif Representation

The PSO K-Means Biclustering Model is chosen to improve the cluster quality which extracts the motif information efficiently without converging to the local minima. The Figure 12 and Figure 13 are the sequence logo representation of the cluster 1 and 2. Similarly the Figure 14 and Figure 15 are the logo representation of the bicluster 1 and 2.

The logo representations of the clusters and biclusters show that the biclusters are more conserved to specific amino acids than the clusters. The horizontal axis shows the positions and the vertical axis shows the bits of amino acids. The PSO K-Means-Biclustering Model is chosen to improve the cluster quality

Table 3. Comparison between clustering and biclustering based on secondary structural similarity

Number of clusters= 147 *Number of biclusters=147 (49× 3)*		
Structural Similarity Value Above	**Clusters**	**Biclusters**
75%	23	35
70%	40	54
65%	69	75
60%	76	89

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Comparison between clusters and biclusters based on structural similarity

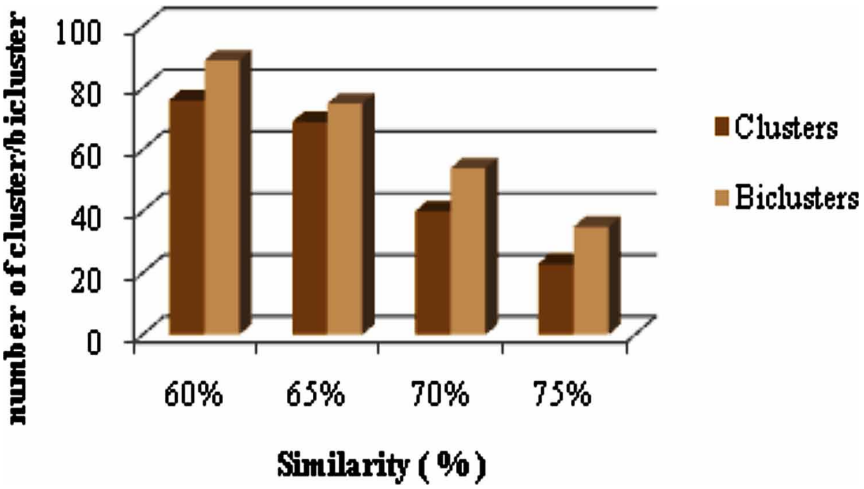


Figure 12. Cluster 1 SAA logo representation

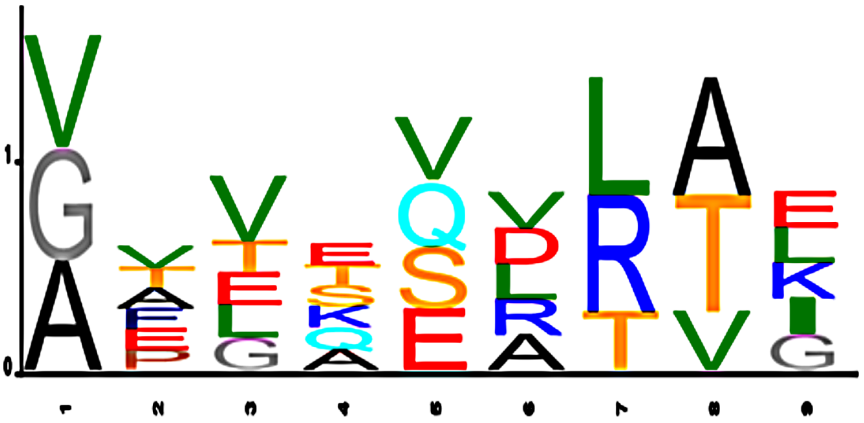


Figure 13. Cluster 2 SAA logo representation

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Bicluster 1 SAA logo representation

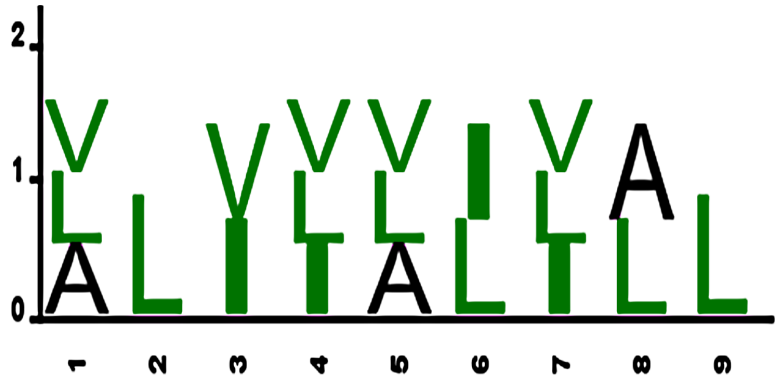


Figure 15. Bicluster 2 SAA logo representation

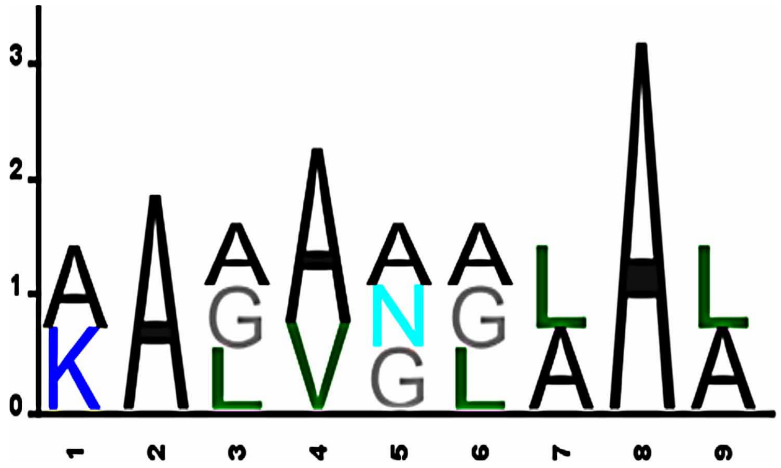


Table 4. Full and partial superset of protein sequence motif on Bicluster 1

Position	SAA	Motif	SAA ⊑ Motif
1	ALV	ADEGILKTV	Full
2	L	ADEGILKTV	Full
3	VI	ADEGILKTV	Full
4	VLI	ADEGILKTV	Full
5	ALV	ADEGILKTV	Full
6	LI	ADEGILKTV	Full
7	VLI	ADEGILKTV	Full
8	AL	ADEGILKTV	Full
9	L	ADEGILKTV	Full

Table 5. Full and partial superset of protein sequence motif on Bicluster 2

Position	SAA	Motif	SAA ⊑ Motif
1	AK	ADEGILKTV	Full
2	A	ADEGILKTV	Full
3	AGL	ADEGILKTV	Full
4	AV	ADEGILKTV	Full
5	ANG	ADEGILKTV	Partial
6	AGL	ADEGILKTV	Full
7	LA	ADEGILKTV	Full
8	A	ADEGILKTV	Full
9	LA	ADEGILKTV	Full

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

because it clusters the columns of the data matrix in addition to clustering the rows. The corresponding amino acid sequence columns that are purged by the PSO K-MeansBiclustering Model (Motif) (Vincent, Bernard, & SinanKockara, n. d.) are shown to be either the full or the partial superset of the significant amino acids (SAA): amino acids with appearing frequency greater than 8%. Table 4 and Table 5 illustrate the SAA and Motif relationship for the entire nine positions of the bicluster 1 and 2. Thus, biclustering are proven to have the potential to select meaningful amino acids that biologists are interested in.

CONCLUSION

The motif information is extracted using this PSO-K-Means model efficiently than simple K-Means. The results show that the PSO-K-Means Biclustering Model is capable of preserving more original data points and obtaining better cluster quality. The biclustering model can detect the highly homologous sequences than the clustering model. The motifs extracted are helpful for biologist for their further research works on new protein sequence. This will be an apt computational process for the biologists handling the two dimensional sequences. It reduces the computational cost when compared to the computations on multi-dimensional sequence analysis.

The future enhancements will be performed as follows:

- Select different number of residues
- Use any criteria for selecting the number of clusters
- Combine granular computing technique with this technique further
- Use some other optimization techniques like BCO, ACO etc.,
- Prediction of biological functionality of the extracted motifs

REFERENCES

Ahmadyfard, A., & Modares, H. (2008). Combining PSO and K-Means to enhance data clustering. *Proceedings of the International Symposium on Telecommunication* (pp. 688 – 691). 10.1109/IS-TEL.2008.4651388

AshaGowdaKaregowda, & SeemaKumari. (2013). Particle Swarm Optimization Algorithm Based K-Means and Fuzzy c-means clustering. *International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(7).

B.Chen, S.Pellicer, P.C.Tai, R.Harrison, & Y.Pan. (2009). Novel efficient granular computing models for protein sequence motifs and structure information discovery. *International Journal of Computational Biology and Drug Design*.

Cheng, B., Hudson, C., Kim, M., Crawford, A., Wright, J., & Che, D. (2011). Protein Sequence Motif Extraction using Decision Forest. Proceedings of the *World Congress in Computer Science, Computer Engineering*.

Bairoch, A., Bucher, P., & Hofmann, K. (1996). The PROSITE database, its status in 1995. *Nucleic Acids Research*, 24(1), 189-196.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bill, C. H., & Chang, S. (2012). Protein Motif Extraction Using Neuro-Fuzzy optimization. *Bioinformatics (Oxford, England)*, *18*(8), 1084–1090. PMID:12176831

Chen, Y.-C., KripamoyAguan, Yang, C.-W., Wang, Y.-T., Pal, N. R., & Chung, I.-F. (2011). Discovery of Protein Phosphorylation Motifs through Exploratory Data Analysis. Discovery of Protein Phosphorylation Motifs(PLoS ONE).

Cui, X., & Potok, T. E. (2005). Document Clustering Analysis Based on Hybrid PSO K-means Algorithm. *Journal of Computer Sciences*, 2005, 27-33.

Das, S., & Idicula, S. M. (2010). Greedy Search-Binary PSO Hybrid for Biclustering Gene Expression Data. *International Journal of Computers and Applications*, *2*(3), 1–5. doi:10.5120/651-908

DEIB. (n. d.). K Means. Retrieved from: www.home.deib.polimi.it/matteucc/clustering/tutorial_html/k_means.html

Elayaraja, E., Thangavel, K., Chitralega, M., & Chandrasekhar, T. (2012). Extraction of Motif Patterns from Protein Sequences Using SVD with Rough K-Means Algorithm. *International Journal of Computer Science Issues*, *9*(6), 350–356.

Elayaraja, E., Thangavel, K., Ramya, B., & Chitralegha, M. (2011). Extraction of Motif Patterns from Protein Sequence using Rough K-Means Algorithm. *Procedia Engineering*, *30*, 814–820. doi:10.1016/j.proeng.2012.01.932

Thangavel, K., Bagyamani, J., & Rathipriya, R. (2011). Novel Hybrid PSO-SA Model for Biclustering of Expression Data. *Proceedings of the International Conference on Communication Technology and System Design* (pp. 1048 – 1055).

Youcheng, L., Nan, T., Majie, S., Kedi, F., Yuan, D., Lincong, Q., et al. (2012). K-means optimization clustering algorithm based on Particle swarm optimization and multiclass merging. In *Advances in CSIE* (Vol. 1, pp. 569-578).

Madeira, S. C., & Oliveira, A. L. (2004). Biclustering Algorithms for Biological Data Analysis. *Survey (London, England)*, 2004, 1–31. PMID:17048406

Riccardo, P., James, K., & Tim, B. (2007). Particle swarm optimization An Overview. In *Swarm Intell.* (pp. 33-57).

Berkhin. (2002). Survey of Clustering Data Mining Techniques. *Accrue Software*.

Vora. (2013). A Survey on K-mean Clustering and Particle Swarm Optimization. *International Journal of Science and Modern Engineering*, 1(3), 24-26.

Rathipriya, R., Thangavel, K., & Bagyamani, J. (2011). Binary Particle Swarm Optimization based Biclustering of Web usage Data. *International Journal of Computers and Applications*, *25*(2), 43–49. doi:10.5120/3001-4036

Pietrokovski, S., Henikoff, J. G., & Henikoff, S. (1996). The BLOCKS database - a system for protein classification. *Nucleic Acids Research*, *24*(1), 197–200. doi:10.1093/nar/24.1.197 PMID:8594578

Reinhard, S., Antoine, D., & Chris, S. (1996). The HSSP database of protein structure–sequence alignments. *Nucleic Acids Research*, 1996, 226–230. PMID:9016541

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Swarm Intelligence. (2006). Particle Swarm Optimization. Retrieved from www.swarmintelligence.org

Kunik, V., Solan, Z., Edelman, S., Ruppin, E., & Horn, D. (2005). Motif Extraction and Protein Classification. proceedings of IEEE. *Computational Systems Bioinformatics / Life Sciences Society. Computational Systems Bioinformatics Conference*.

Wang, G., & Dunbrack, R. (2003). PISCES: a protein sequence culling server in Bioinformatics, 19(12), 1589-1591.

Cluster Analysis. (n. d.) *Wikipedia*. Retrieved from: en.wikipedia.org/wiki/Cluster_analysis

Fujiwara, Y., & Konagaya, A. (2008). Protein Motif Extraction using Hidden Markov Model. Proceedings of the Genome Informatics Workshop IV (pp. 57-64).

Vincent, Y., Bernard, C., & Kockara, S. (n. d.). Extraction of Protein Sequence Motifs Information by Bi-Clustering Algorithm.

Zhaoa-Xing-Ming. (2005). A novel approach to extracting features from motif content and protein composition for protein sequence classification. Neural Networks, 2005, 1019–1028. PubMed PMID:16153801

KEY TERMS AND DEFINITIONS

Biclustering: The process of grouping data based on both the rows (segments) and columns (amino acids)

Bio-Inspired Computing: Bio-inspired computing is a field devoted to tackling complex problems using computational methods modeled after design principles encountered in nature

Clustering: Grouping similar kind of data elements. It is used to discover similar patterns from a sea of data. The similarity between the objects in the same cluster is greater than that of the different clusters.

Protein Motif: Motifs are the patterns that present repeatedly in the sequence which are responsible for various cellular processes

Protein Sequence: Sequence of Amino Acids responsible for various cellular activities of the biological systems

PSO: Particle Swarm Optimization is one of the swarm based optimization technique. It is devised based on the bird flocking behavior

Significant Amino Acids (SAA): SAA represents the protein pattern that present in the set of protein sequences that shows the dominant characteristic present in that Protein group.

ENDNOTE

[1] Structures of life (Book)

This research was previously published in the Handbook of Research on Computational Intelligence Applications in Bioinformatics edited by Sujata Dash and Bidyadhar Subudhi; pages 240-262; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 66
Classification of Breast Thermograms Using Statistical Moments and Entropy Features with Probabilistic Neural Networks

Natarajan Sriraam
M. S. Ramaiah Institute of Technology, India

Prabha Ravi
M. S. Ramaiah Institute of Technology, India

Leema Murali
M. S. Ramaiah Institute of Technology, India

B. Venkataraman
Indira Gandhi Centre for Atomic Research, India

Amoolya Girish
M. S. Ramaiah Institute of Technology, India

M. Menaka
Indira Gandhi Centre for Atomic Research, India

Manjunath Sirur
M. S. Ramaiah Institute of Technology, India

A. Shenbagavalli
National Engineering College, India

Sushmitha Srinivas
M. S. Ramaiah Institute of Technology, India

Josephine Jeyanathan
Kalasalingam University, India

ABSTRACT

Breast cancer is considered as one of the life-threatening disease among woman population in developing as well as developed countries. This specific study reports on classification of breast thermograms using probabilistic neural network (PNN) with four statistical moments features mean, standard deviation, skewness and kurtosis and two entropy features, Shannon entropy and Wavelet packet entropy. The CLAHE histogram equalization algorithm with uniform and Rayleigh distributions were considered for contrast enhancement of breast thermal images. The asymmetry detection was performed by applying bilateral ratio. A total of 95 test images (normal = 53, abnormal = 42) was considered. Simulation

DOI: 10.4018/978-1-7998-1204-3.ch066

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

study shows that CLAHE -RD with wavelet entropy features confirms the existence of symmetry on the right and left breast thermal images. An overall classification accuracy of 92.5% was achieved using the proposed multifeatures with PNN classifier. The proposed technique thus confirms the suitability as a screening tool for asymmetry detection as well as classification of breast thermograms.

INTRODUCTION

Medical thermography is found to be a promising tool for early breast cancer detection and provides a scope of developing computer aided diagnostics tool for the biomedical engineering community. The high metabolic changes are reflected in the form of high temperature in the localized region through infrared thermal images and this property is well exploited for the clinical breast cancer interpretation (Kontos et al., 2011; Amri et al., 2016). The inherent relation between the radiation from the human skin and surface temperature has a huge influence in the level of skin's blood perfusion. This property is being exploited for infrared thermal imaging and the effect of angiogenesis, inflammatory, etc. is well captured by this imaging procedure. It has the ability to showcase the changes in the vascular process referred as neoangiogenesis (Nishida et al., 2006). The increase in surface temperature in the cancerous regions can be well exploited using this imaging procedure. The active and inactive blood vessels in and around the breast region can be well mapped through this thermographic imaging. It further detects the newly formed or activated blood vessels with distinct appearance. Due to its ability to detect changes at the cellular level, study reveals that the thermographic imaging test can detect activity 8 to 10 years before any other test. This makes the IR imaging based thermography as the potential candidate to detect the metabolic changes before the actual formation of the tumor (Moghbel et al., 2011; Bhowmik et al., 2015; Amri et al., 2016; Garduño-Ramón et al., 2017).

Several attempts have been made towards the detection of breast cancer using thermography imaging technique. Kuruganti & Qi (2002) performed asymmetric analysis in breast cancer detection. The hotspot cancerous tissue due to temperature variation was recognized by this approach. Initial segmentation was performed using Hough transform. The test images were obtained from Bioyear Inc. with a sensitivity of 0.05 degrees Celsius.

Mahmoud Zadeh et al. (2015) have applied the extended hidden Markov model for optimal segmentation of breast thermal patterns. The proposed method was able to recognize semi hot regions into distinct areas. The computational burden was less compared to other segmentation procedures reported in the literature. A specific study on breast cancer detection using rotational thermography was reported (Francis et al., 2014). The proposed work showed a significant improvement on detection of breast cancer compared to the traditional thermography procedure. Textural features were extracted to distinguish cancer and non-cancerous tissues from breast thermo grams. Acharya et al. (2010) have showed the effect of textural features with support vector machine classifier for breast cancer detection using infrared thermal images. An overall classification accuracy of 88.10% was attained based on the database obtained from General hospital, Singapore. Bhowmik et al. (2015) have shown a specific study on analysis of hybrid intelligent techniques using breast thermography the detection of breast cancer. A specific study by making use of higher order statistics for breast thermograms was reported (Acharya et al., 2014). Ramakrishnan & Suganthi have investigated the effect of Gabor wavelet transform for breast thermogram images (Ramakrishnan & Suganthi, 2014).

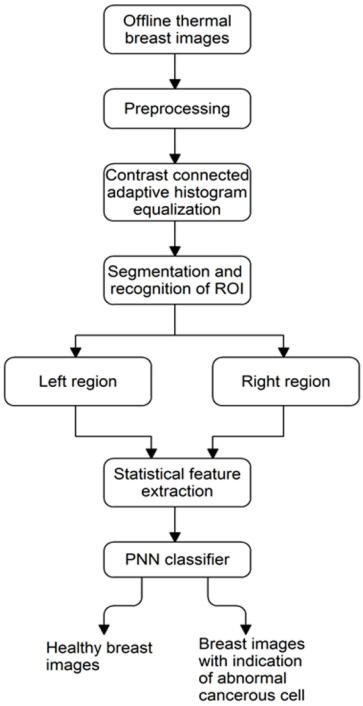
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This specific research study investigates the classification of breast thermograms for cancer detection using the thermal breast images collected from Indian women ethnic group. The pilot study was carried out to develop automated screening algorithm for breast cancer among Indian woman. The proposed scheme employs CLAHE with Rayleigh distribution for contrast enhancement followed by extraction of statistical moments and entropy features. The semi-automated region of interest detection helps in identifying the symmetry/asymmetry nature of the acquired breast thermograms. Later probabilistic neural network was applied for classifying the normal and abnormal breast thermograms. Ductal carcinoma was considered as the abnormal for the proposed study. Further the study was restricted to lateral position of the thermograms for the investigations.

PROPOSED BREAST THERMOGRAM CLASSIFICATION SCHEME

Figure 1 shows the flow diagram of the proposed breast thermogram classification technique. The breast thermal images were captured using the FLIR IR camera using the standard protocol with appropriate ethical clearance. The acquired offline images were filtered and grayscale image processing was performed. In order to enhance the image, Contrast-Limited Adaptive Histogram Equalization (CLAHE) mapping to two variants, uniform and Rayleigh distributions was applied. Canny edge detection with gradient operator was introduced to segment the image and appropriate ROI was selected through semi-automated technique from right and left breast regions.

Figure 1. Proposed breast thermogram classification scheme



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Four statistical moments such as, mean, standard deviation, skewness and kurtosis along with two entropies, Shannon's entropy and Wavelet packet entropy were applied and these significant features later was collectively fed to a probabilistic neural network model for classification of normal and abnormal breast images. In order to locate the existence of asymmetry between the left and right breast regions, a parameter referred as bilateral ratio was applied. The performance of the proposed model was evaluated in terms of sensitivity, specificity and classification accuracy.

STUDY SETUP

Figure 2 shows the screening setup of breast thermography. The subjects were requested to sit in a comfortable position and were asked to undress up to waist to ensure smooth recording of the image under controlled room temperature (22° C). Due ethical clearance was obtained and subject consent form were collected before the commencement of the procedure. The camera was arranged at a distance of 1.5m from the subject. Only frontal postures were considered for the analysis in this pilot study. The database was based on the breast thermograms collected from Indian ethnic women of age group 24-58.

Being a non-invasive screening technique, breast thermogram images cannot be compared directly to recognize the normal and cancerous cells. The asymmetry analysis indicates the presence of cancerous cells. It is well known fact that the red and white regions in the breast thermograms indicate the hot spots; green region as moderate temperature and blue regions as cold regions. The temperature distribution plays a very crucial role in breast region analysis. Figures 3 and 4 show the typical normal and abnormal images. For this specific study, 53 normal and 42 abnormal breast thermal images were collected.

GRAY SCALE PROCESSING

The thermal images were filtered using second order Butterworth filter and gray scale level processing was applied. In order to improve the contrast of the image, a variant of histogram equalization referred as contrast Limited Adaptive Histogram Equalization (CLAHE) was applied. To select the best region of interest, two different mapping was considered: Uniform and Rayleigh distributions (Erturk, &Konyar.,

Figure 2. Screening setup of breast thermography

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Typical normal images

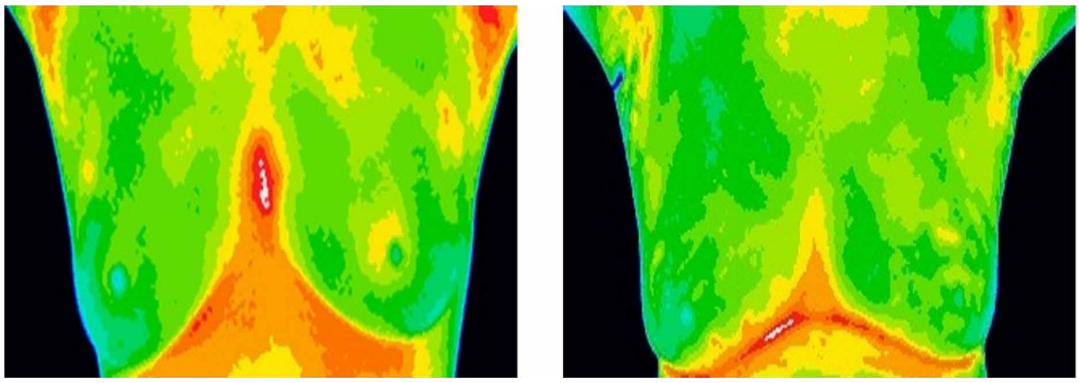
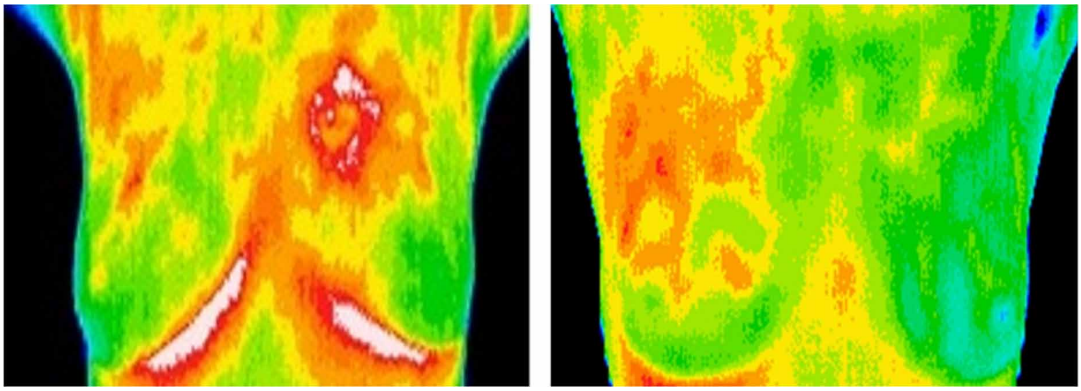


Figure 4. Typical abnormal images (ductal carcinoma)



2015; Erturk, 2013). While considering the infrared images, there are certain cases where uniform re-distribution suffers due to nondiscrimination of background and foreground and allows noise to some extent (Yousefi et al., 2013). This study compares the influence of both distributions. Each histogram in the CLAHE refers to distinct section of the image referred to as tiles (Pizer et al., 1990). The tiles contrast enhanced the results in the specific distribution. Figure 5 (parts a-e) shows the entire process.

SEGMENTATION OF ROI

Segmentation of ROI is an important step in image processing, it is from this image that the accuracy of the system is determined. Segmentation follows pre-processing which includes background removal and resizing for removing the undesired body portions in the breast thermogram images. The resultant image is then subjected to canny edge detection to extract the breast boundaries, suitable threshold value is chosen to eliminate detection of extra curves (Scales et al., 2004). The next step includes the extraction of left and right boundary curves using gradient operator. The lower breast boundaries are extracted and the unwanted portions below it are removed and separation line is created to separate the right and left regions. Figure 6 depicts the resultant ROI with left and right breast regions.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. (a) Sample image; (b) Gray scale image; (c) Histogram; (d) Contrast enhanced image; (e) Histogram

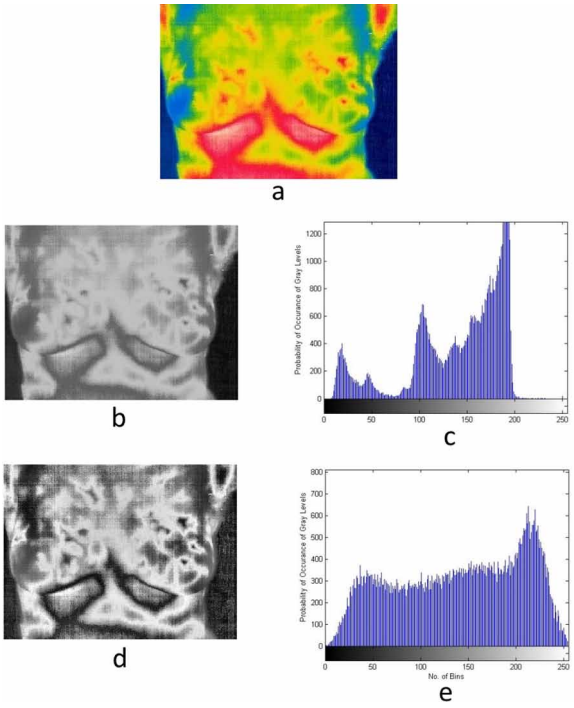
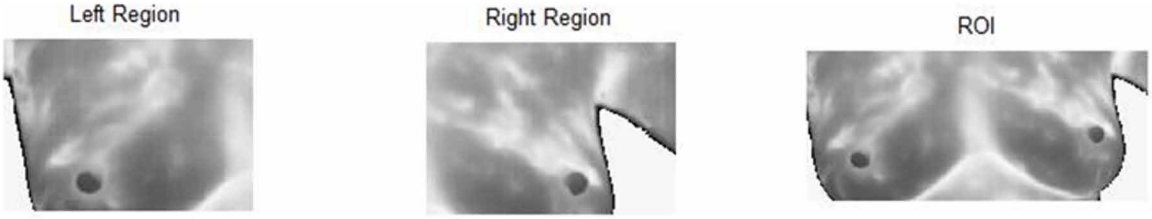


Figure 6. ROI with left and right segmented portions



FEATURE EXTRACTION

After the thermal images were preprocessed, the resultant image was dimensionally segmented into two regions namely right and left. The following statistical moments were applied for the proposed work (Qi & Kuruganti, 2002; Ng et al., 2001):

1. **Mean:** It is average value of pixels in the region of interest. Slight variation of mean is seen in noncancerous patients and vice versa in cancerous patients. It is calculated by:

$$\mu = 1 / MN \sum_{i=1}^{M} \sum_{j=1}^{N} \left(P(i,j) - \mu \right)^2$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. **Standard deviation:** It is a measure of quantifiable amount of variance. It also measures pixels distribution. Once the preprocessed image is dimensionally segmented into right and left halves, above mentioned statistical features are extracted from both the regions. Standard deviation is calculated by taking the square root of variance which is given by:

$$\sigma = \sqrt{\mu = 1 / MN \sum_{i=1}^{M} \sum_{j=1}^{N} \left(P(i,j) - \mu \right)^2}$$

3. **Kurtosis:** It identifies the abnormality and classifies as normal or abnormal class. The image with high kurtosis have shaper peak near mean and the image with low kurtosis have flat peak near the mean. Kurtosis is calculated by:

$$K = \left\{ 1 / MN \sum_{i=1}^{M} \sum_{j=1}^{N} \left(\frac{P(i,j) - \mu}{\sigma} \right)^4 - 3 \right\}$$

4. **Skewness:** Measures asymmetry of the pixel value around the sample mean. Negative skewness indicates that the pixels are spread out more towards left side of the mean. Positive skewness indicates that the pixels are spread out more towards right side of the mean. Ideally skewness of normal distribution is zero. Skewness is calculated by:

$$S = 1 / MN \sum_{i=1}^{M} \sum_{j=1}^{N} \left(\frac{P(i,j) - \mu}{\sigma} \right)^3$$

Two entropies, Shannon's Entropy (SE) and Wavelet Packets Entropy (WE)were applied as features apart from the statistic features. The Wavelet Packet Entropy (WE) was calculated based on the work reported by Wong and Cheng (2013).

Figure 7 shows the plot of four statistical moments based on CLAHE with Uniform Distribution (UD) and Rayleigh Distribution (RD) respectively for right and left breast regions.

ASYMMETRY IDENTIFICATION

In order to recognize whether the test breast thermal image has cancerous tissue or not, asymmetry test was performed using the parameter, Bilateral Ratio (Kuruganti & Qi, 2004). It can be estimated based on the feature values calculated for right and left breast regions. It is understood that closer the value to one, more existence of correlation among the left and right regions that confirms the symmetry behavior. This aspect indicates that the given test image can be categorized as "normal." The Bilateral Ratio (BR) (closeness to 1) is defined as:

$$\mathbf{BR} = \left| \frac{\textbf{Feature value from left}}{\textbf{Feature value from right}} - \mathbf{1} \right|$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Feature plot distribution based on CLAHE-RD and CLAHE-UD

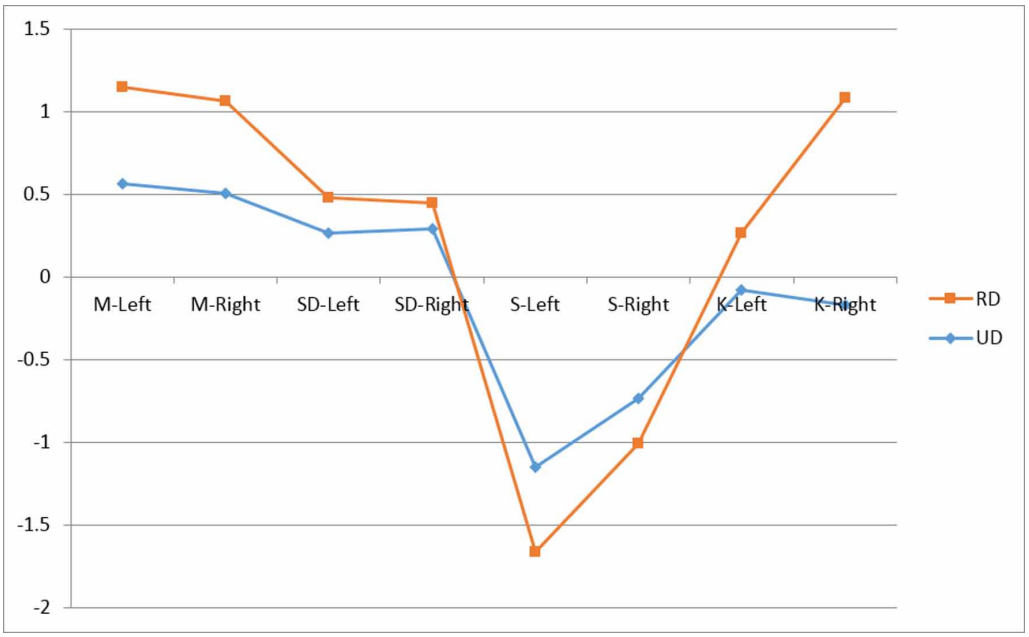


Figure 8 shows the BR results for statistical and entropy features. The influence of CLAHE with histogram distribution was also studied.

One can infer from Figure 8 that the Rayleigh's Distribution outperforms with Uniform Distribution. Among the four moments, entropy features, Wavelet Packet Entropy confirms its suitability with Rayleigh's Distribution for Asymmetry identification between left and right breast regions.

PERFORMANCE EVALUATION

The performance of the proposed scheme was evaluated by applying probabilistic neural network (PNN) model as classifier. PNN is a basically a radial/basis neural network which works on the principle of probabilistic theory typically a Bayesian model. It comprises of input layer, pattern layer, summation layer and an output layer. For the proposed scheme, 6 input nodes were used. One can refer Morey et al., 2016 for further functioning of PNN model. The spread factor was fixed as 0.08. The PNN classifier was evaluated in terms of sensitivity, specificity and classification accuracy. A three-fold cross validation procedure was adopted using all the given features. The classification accuracy was estimated based on (6):

$$\mathbf{CA} = \frac{\mathbf{No.\ of\ Correctly\ Detected\ Breast\ Images}}{\mathbf{Total\ no.\ of\ applied\ Breast\ Images}}$$

Figures 9 and 10 show the training and testing results obtained using PNN classifier.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Plot of estimating bilateral ratio

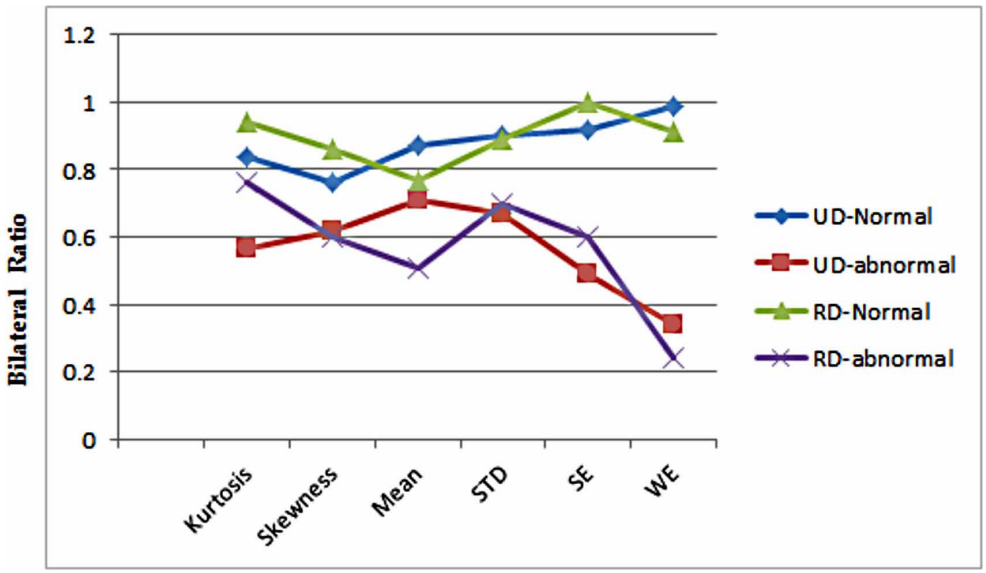
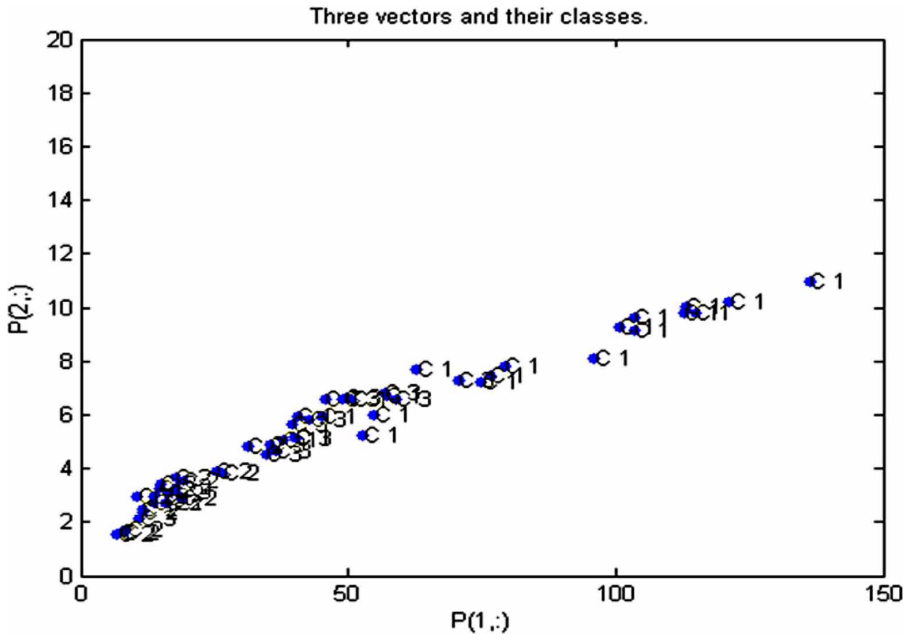


Figure 9. Three vectors – Training (PNN)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11 shows the classification results using UD and RD distribution. It can be seen that the PNN classifier with multi features extracted based CLAHE-RD outperforms then CLAHE-UD. An overall CA of 92.5% was achieved.

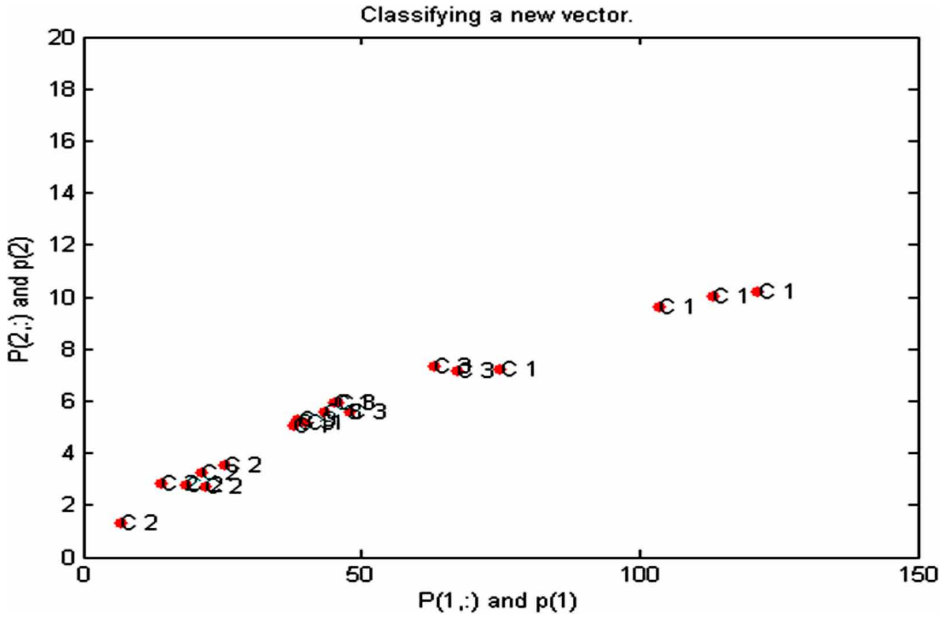
The classifier performance was further analyzed using the receiver operating characteristics (ROC) and result is shown in Figure 12.

The area under curve was found to 93.5% and it can be seen the ROC study confirms the suitability of the proposed pattern classifier with the selected multifeatures.

CONCLUDING REMARKS

Automated screening of breast thermograms using statistical and entropy features with probabilistic neural network as classifier was discussed in the proposed study. A set of normal (n=53) and abnormal with indication of cancerous tissue (n = 42) was considered from a specific ethnic group. The contract enhancement procedure was done using CLAHE with uniform and Rayleigh's distribution. The ROI was selected based on canny edge detection followed by thresholding and the given image was segmented into left and right regions. The asymmetry behavior was tested using the Bilateral ratio. The entropy features with CLAHE-RD was found to be efficient in recognizing the level of asymmetry between the right and the left breast regions. All the six features extracted from the selected ROI were applied to the PNN classifier. The three-fold cross validation showed a classification accuracy of 92.5%. The current procedure can be introduced for screening of breast cancer on the woman population. The effect of CLAHE-RD and PNN classifier with the applied features can be validated with a large dataset for its clinical suitability.

Figure 10. Classification of new test set



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Performance of PNN classifier

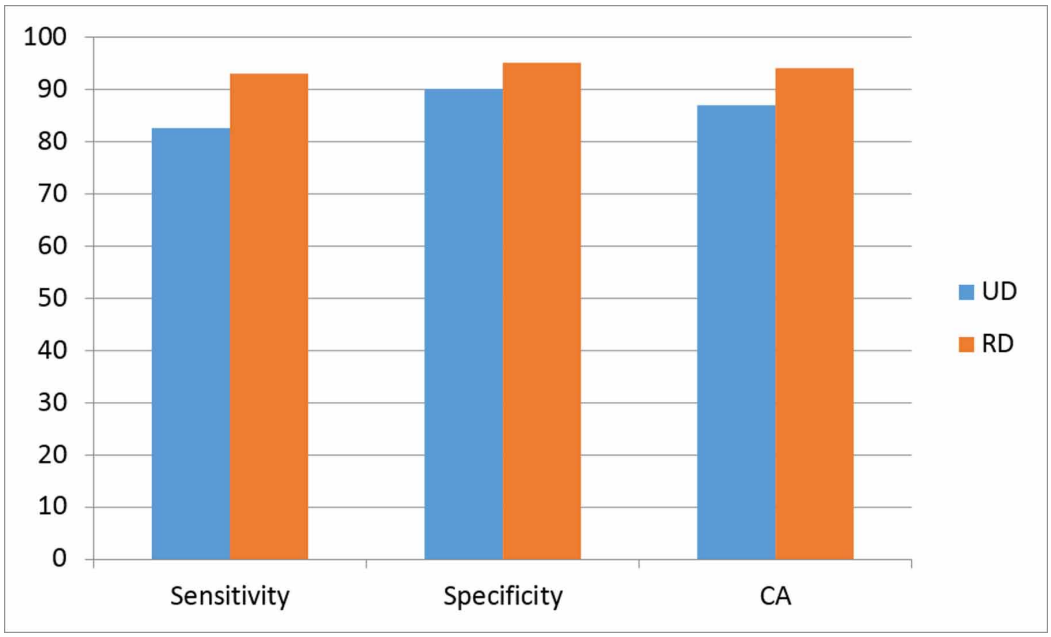
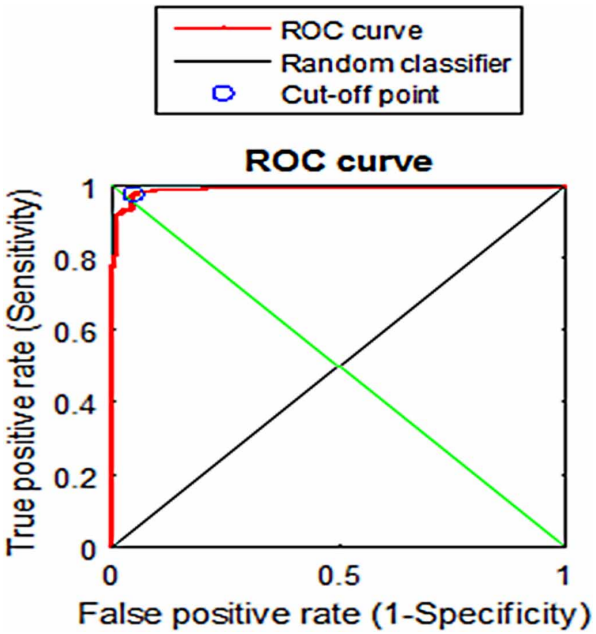


Figure 12. ROC curve of PNN classifier



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Acharya, U. R., Chua, C. K., Chattopadhyay, S., Ng, E. Y., & Sree, S. V. (2014). Higher order spectra analysis of breast thermograms for the automated identification of breast cancer. *Expert Systems: International Journal of Knowledge Engineering and Neural Networks*, *31*(1), 37–47. doi:10.1111/j.1468-0394.2012.00654.x

Acharya, U. R., Ng, E. Y., Sree, S. V., & Tan, J. (2010). Thermography based breast cancer detection using texture features and support vector machine. *Journal of Medical Systems*, *36*(3), 1503–1510. doi:10.100710916-010-9611-z PMID:20957511

Amri, A., Pulko, S. H., & Wilkinson, A. J. (2016). Potentialities of steady-state and transient thermography in breast tumour depth detection: A numerical study. *Computer Methods and Programs in Biomedicine*, *123*, 68–80. doi:10.1016/j.cmpb.2015.09.014 PMID:26522612

Bhowmik, M. K., Bhattacharjee, D., Gogoi, U. R., Ghosh, A. K., & Majumdar, G. (2015). Breast abnormality detection through statistical feature analysis using infrared thermograms. In *Proceedings of the 2015 International Symposium on Advanced Computing and Communication (ISACC)* (pp. 258-265).

Erturk, S. (2013). Improved region of interest for infrared images using Rayleigh Contrast-Limited Adaptive Histogram Equalization. Retrieved from http://worldcomp-proceedings.com/proc/p2013/IPC2477.pdf

Erturk, S., & Konyar, M. Z. (2015). Enhancement of ultrasound images with bilateral filter and Rayleigh CLAHE. In *Proceedings of the 2015 23nd Signal Processing and Communications Applications Conference (SIU)* (pp. 1861-1864).

Francis, S. V., Sasikala, M., & Saranya, S. (2014). Detection of Breast Abnormality from Thermograms Using Curvelet Transform Based Feature Extraction. *Journal of Medical Systems*, *38*(4), 1–9. doi:10.100710916-014-0023-3 PMID:24659445

Garduño-Ramón, M. A., Morales-Hernández, L. A., Osornio-Rios, R. A., & Vega-Mancilla, S. G. (2017). Supportive noninvasive tool for the diagnosis of breast cancer using a thermographic camera as sensor. *Sensors (Basel, Switzerland)*. PMID:28273793

Kontos M, Wilson R, Fentiman I. (2011). Digital infrared thermal imaging (DITI) of breast lesions: sensitivity and specificity of detection of primary breast cancers. *Clin. Radiol.*, *66*(6), 536-539.

Kuruganti, P.T., & Qi, H. (2004). Detecting breast cancer from thermal infrared images by asymmetry analysis.

Mahmoudzadeh, E., Montazeri, M. A., Zekri, M., & Sadri, S. (2015). Extended hidden Markov model for optimized segmentation of breast thermography images. *Infrared Physics & Technology*, *72*, 19–28. doi:10.1016/j.infrared.2015.06.012

Moghbel, M., & Mashohor, S. (2011). A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection. *Artificial Intelligence Review*, *39*(4), 305–313. doi:10.100710462-011-9274-2

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Morey, P., Nagtode, S., & Potdukhe, B. B. (2016). Two dimensional discrete Wavelet transform and Probabilistic neural network used for brain tumor detection and classification. In *Proceedings of the 2016 Fifth International Conference on Eco-friendly Computing and Communication Systems (ICECCS)* (pp. 20-26).

Ng, E.Y.K., Ung, L.N., Ng, F.C., & Sim, E. (2001). Statistical analysis of healthy and malignant breast thermography. Journal of Medical Engineering & Technology, 25(6), 253–263. doi:10.1080/03091900110086642 PMID:11780767

Nishida, N., Yano, H., Nishida, T., Kamura, T., & Kojiro, M. (2006). Angiogenesis in Cancer. *Vascular Health and Risk Management*, *2*(3), 213–219. doi:10.2147/vhrm.2006.2.3.213 PMID:17326328

Pizer, S. M., Johnston, R. E., Ericksen, J. P., Yankaskas, B. C., & Muller, K. E. (1990). Contrast-limited adaptive histogram equalization: speed and effectiveness. In *Proceedings of the First Conference on Visualization in Biomedical Computing* (pp. 337-345). 10.1109/VBC.1990.109340

Qi, H., & Kuruganti, P. T. 2002. Asymmetry Analysis in Breast Cancer Detection Using Thermal Infrared Images. In *Proceedings of the Second Joint EMBSiBMES Conference*, Houston, TX, October 23-26. IEEE

Ramakrishnan, S., & Suganthi, S. S. (2014). Analysis of Breast Thermograms Using Gabor Wavelet Anisotropy Index. *Journal of Medical Systems*, *38*, 1–7. PMID:24395031

Scales, N., Herry, C., & Frize, M. (2004). Automated image segmentation for breast analysis using infrared images. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society Conference* (Vol. *3*, pp. 1737–1740). PMID:17272041

Wong, T. K., & Cheng, C. H. (2013). Automated Segmentation and Hybrid Classifier for Identifying Medical Image. International Journal of Signal Processing. *Image Processing and Pattern Recognition*, *6*(1), 191–202.

Yousefi, S., Qin, J., Zhi, Z., & Wang, R. K. (2013). Uniform enhancement of optical micro-angiography images using Rayleigh contrast-limited adaptive histogram equalization. *Quantitative Imaging in Medicine and Surgery*, *3*(1), 5–17. doi:10.3978/j.issn.2223-4292.2013.01.01 PMID:23482880

This research was previously published in the International Journal of Biomedical and Clinical Engineering (IJBCE), 6(2); edited by Natarajan Sriraam; pages 18-32; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 67
HIPAA Security and Privacy Rules Auditing in Extreme Programming Environments

Mahmood Alsaadi
Princess Sumaya University for Technology, Jordan

Malik Qasaimeh
Princess Sumaya University for Technology, Jordan

Sara Tedmori
Princess Sumaya University for Technology, Jordan

Khaled Almakadmeh
Hashemite University, Jordan

ABSTRACT

Healthcare business is responsible of keeping patient data safe and secure by following the rules of the federal Health Insurance Portability and Accountability Act of 1996, (HIPAA). Agile software organizations that deal with healthcare software system face a number of challenges to demonstrate that their process activities conform to the rules of HIPAA. Such organizations must establish a software process life cycle and develop procedures, tools, and methodologies that can manage the HIPAA requirements during the different stages of system development, and also must provide evidences of HIPAA conformity. This paper proposes an auditing model for HIPAA security and privacy rules in XP environments. The design of the proposed model is based on an evaluation theory which takes as its input the work of Lopez ATAM, and the standards of common criteria (CC) concepts. The proposed auditing model has been assessed based on four case studies. The auditing result shows that the proposed model is capable of capturing the auditing evidences in most of the selected case studies.

DOI: 10.4018/978-1-7998-1204-3.ch067

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

In a world of increasingly global competition, the competitive demands on companies and organizations in the healthcare industry become more intense (Wall, 2009). This has induced healthcare organisations to exploit new opportunities to gain competitive advantage (Reina, Lacroce, Cetani, & Ventura, 2012).

Regulatory compliance has become visible in healthcare industries. The federal Health Insurance Probability and Accountability Act of 1996 (HIPAA), passed by the United States Congress and signed by President Bill Clinton, is the first comprehensive federal guideline for the privacy of patients and health information (Brian & Daniel, 1997). HIPAA is a set of rules aimed at strengthening patients' rights, whilst decreasing administrative costs in the healthcare industry. Failure to comply with HIPAA could lead to large penalties and in extreme cases could lead to loss of medical licenses. Organizations that deal with protected health information (PHI) and wishing to be HIPAA certificated must follow the rules of HIPAA (i.e. physical requirements, network requirements, and security and privacy requirements).

HIPAA is designed to: 1) improve the quality of health insurance; 2) improve the portability, and continuity of health insurance coverage in the group and individual market; 3) simplify the administration of health insurance; 4) prevent fraud and corruption in health insurance and health care companies; 5) protect a subset of sensitive information known as protected health information (PHI); and 6) protect health data created, received, maintained or transmitted electronically, also known as electronic protected health information (ePHI) (Brian & Daniel, 1997).

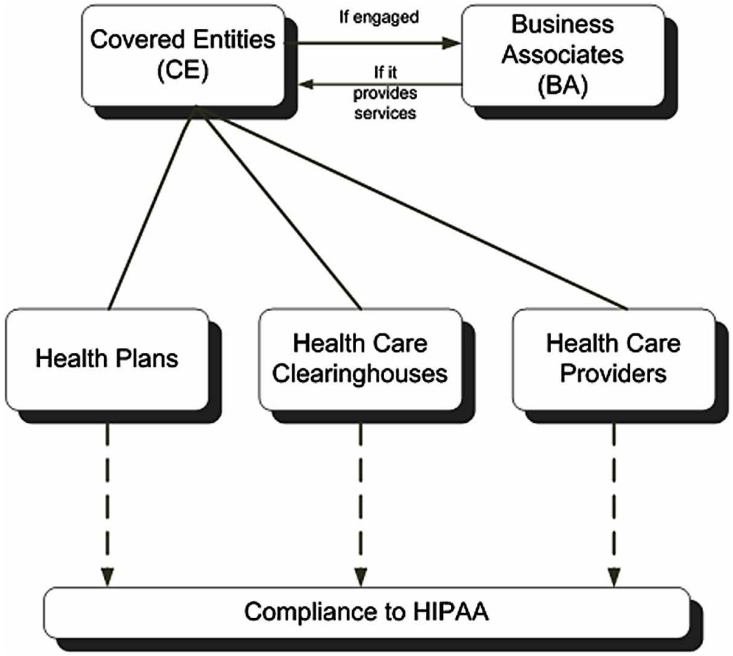
Organizations that deal with health information must comply with the HIPAA rules. In general, compliance means conforming to the authoritative rules in order to get certification for specific use (Ahmed, 2014). HIPAA rules apply to both covered entities and business associates. For that, industries or organizations that handle protected health information (PHI) or electronic protected health information (ePHI) are related to any of the covered entities or business associates. Next is a description of entities involved in the HIPAA compliance process, see Figure 1.

- **Covered Entities (CE):** According to HIPAA, the term "covered entity" refers to three specific businesses including: health plans, health care clearinghouses, and health care providers that transmit health information electronically (U.S. Department of Health & Human Services, 2014). Examples include software organizations (third parties), hospitals, pharmacies, academic medical centers, and other health care providers. Covered entities can be institutions, organizations, or persons.
- **Business Associates (BA):** The Department of Health and Human Services (U.S. Department of Health & Human Services, 2014) defined the term business associate as "a person or entity that performs certain functions or activities that involve the use or disclosure of protected health information (PHI) on behalf of, or provides services to, a covered entity". Examples of business associates include individuals who perform services as part of the workforce of a covered entity, financial and banking institutions when performing payment processing activities, medical transcription companies, and others such as: auditors, consulting firms, or software vendors and consultants.

Extreme programming (XP) is an agile software development process that is designed to simplify and expedite the process of developing software products by favouring close collaboration between software development and business teams via face-to-face communication, as opposed to putting an emphasis on written documentation.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Entities of HIPAA Compliance Process



Several studies can be found in literature indicating the high adoption of XP practices when developing medical device software (AAMI, 2012; McHugh, McCaffery, & Casey, 2014; Fitzgerald, Stol, O'Sullivan, & O'Brien, 2013). These studies are evidence of medical software based organizations shifting away from the use of traditional development processes towards agile development processes. Moreover, these studies show how such software originations are recently paying more attention to the compliance process with regulatory requirements.

A review of the research literature on the practices of XP has revealed weaknesses in its capabilities to satisfy the requirements of HIPAA rules. These weaknesses can be attributed to the gaps that exist between the HIPAA rules and the XP software processes such as:

- **Documentation:** HIPAA has strict documentation requirements that industries must comply with. On the other hand, XP aims to keep documentation as simple as possible.
- **Traceability:** HIPAA regulation requires documented evidence on how risk analysis has been done, and how the requirements (security and privacy rules) have been implemented, designed, and tested. While, XP uses simple documentation for traceability.

The aim of this paper is to enhance the auditability of XP in response to HIPAA, as well as to help software organizations that use XP when developing medical software to become HIPAA certified. This paper proposes an auditing model that 1) enables the HIPAA auditors to obtain evidence in conformity to HIPAA security and privacy rules, 2) provides HIPAA auditors with methodological approach of auditing process, and 3) helps the software developers in the medical environment to develop their software system with HIPAA rules in mind.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The design methodology for the proposed auditing model is based on evaluation theory (ATAM) and Common Criteria standards (CC) as well as an evaluation procedure proposed by CC concepts. The proposed auditing model has been assessed based on four case studies to investigate the compliably of the selected case studies with the HIPAA security and privacy rules.

The paper is organized as follows. In addition to the first section which provided an introduction to the research area of HIPAA compliance in medical software based systems, section 2 provides an overview of related works. Section 3 provides an analysis of HIPAA auditing requirements. Section 4 provides a description of XP process. Sections 5 and 6 provide a description of the proposed audit model. Section 7 provides an assessment of the proposed auditing model through a detailed auditing of selected case studies. Section 8 discusses the impact of the proposed auditing model on the XP process. Finally, section 9 concludes the research and discusses avenues for future work.

2. RELATED WORK

McHugh et al. (2013) present a study of agile software techniques adoption in software organizations that develop medical device in Ireland. In their paper, the researchers conducted a comparison between plan-driven software development life cycle and agile software development and then prepared some recommendations on how medical device software development organizations can modify their existing structure to become more suitable for adopting agile development techniques. Moreover, the researchers highlighted that when a medical device software development organization wants to use agile processes, it must fix its current organizational structure based on the use of home-ground analysis which examines five critical success factors for adopting agile methods within software organizations.

McHugh, McCaffery, Casey, and Pikkarainen (2012) discuss the challenges faced by developers when developing medical device software and how the shortcomings of both agile and plan driven approaches can be resolved by following a mixed method approach towards the success of medical device software development. The researchers point out that a single agile methodology is not appropriate to use when developing medical device software. For this reason, the researchers propose a mixed method approach which combines agile practices with plan driven software development lifecycle. Furthermore, the researchers suggest developing a tailored V-Model which integrates the agile practices to facilitate the achievement of regulatory compliance. The researchers do not explain how agile software development can meet the authoritative rules such as HIPAA.

Mehrfard, Pirzadeh, & Hamou-Lhadj (2010) focus on the Food and Drug Administration (FDA) regulations. The researchers discuss FDA requirements in details and show how XP lacks the necessary practices to support FDA. The researchers propose a mapping process between FDA and XP by focusing on human factor engineering requirements that must be met by any software process that claims to be FDA compliant. Moreover, the researchers show how XP does not support FDA. Additionally, the researchers propose an extension to the XP exploration phase which involves several sub-phases such as, contextual investigation, requirements scenario, vision/use scenario and stories. In their work, the researchers did not focus on security requirements and their generic framework is limited to a specific life-science regulation (i.e. FDA).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Qasaimeh and Abran (2011) propose an extension to the user story of XP as to enhance the ability of user stories to collect the information required for achieving ISO 9001 certification. The extension is based on sub processes related to the Capability Maturity Model- Integration (CMMI) model: 1) identification of the source of the user story; 2) categorization of the nonfunctional requirements; 3) identification of the user story relationships; and 4) prioritization of the user stories.

Qasaimeh and Abran (2013) proposed an audit model for ISO 9001 traceability requirements in XP environments. The objective was to help the IS-auditors in the process of extracting the auditing evidences that demonstrate the conformance of XP to ISO 9001 requirements. What distinguishes this paper from Qasaimeh and Abran's (2013) research is that their proposed audit model works only with ISO 9001 which is a general regulation for any software developing organization, but not focuses on specific development requirements. The contribution of this research is specific to HIPAA security and privacy requirements.

3. ANALYSIS OF HIPAA AUDITING REQUIREMENTS

According to HIPAA privacy and security rules, an organization that deals with patient health information must protect the confidentiality of health data for the individuals' by developing certain security controls and monitoring activities that comply with HIPAA security and privacy rule. For example, a security audit must be maintained at different levels of abstraction during the software development life cycle. The HIPAA security and privacy rule includes two provisions that impact the healthcare organizations security audits (HIPAA, 1996). The two extracted provisions are:

Section 64.308(a)(1)(ii)(c), Information system review, which states that organizations must "implement procedures to regularly review records of information system activity, such as audit logs, access reports, and security incident tracking reports". According to the HIPAA auditing rules, the information system review refers to monitoring the use of sensitive information that act as evidence in the process information security incidents investigation. In software engineering, information system review is the process of software product examined by project personnel, managers, users, customers, user representatives, or other interested parties for comment or approval (Wiegers, 2002). Moreover, the information system review is one of the HIPAA auditing procedures to ensure that the software products that deal with healthcare software systems or access the ePHI have been reviewed, audited, verified, to meet the required requirements of HIPAA rules.

Section 164.312(1) (b), Audit controls, which states that organizations must "implement hardware, software, and procedural mechanisms that record and examine activity in information systems that contain or use electronic protected health information". According to the HIPAA rules, audit controls are required in order to enforce the software organization into implementing auditing techniques to record and monitor each activity in any product or information system that implements ePHI. In software engineering, auditing can be divided into two categories (Westfall, 2007):

- **Functional Audit:** This type of audit is used to verify that all requirements specified in the software requirements specification have been met.
- **Physical Audit:** This type of audit is used to verify that the software product and its related documentation are internally consistent.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Audit controls are essential and required by the HIPAA rules to ensure that the requirements specified in the software specification have been met, and that the system operates correctly. While, HIPAA focuses on security and privacy, audit controls can be considered as the process of reviewing, monitoring, capturing, storing, analyzing, and reporting information related to security and privacy threats.

4. XP PROCESS

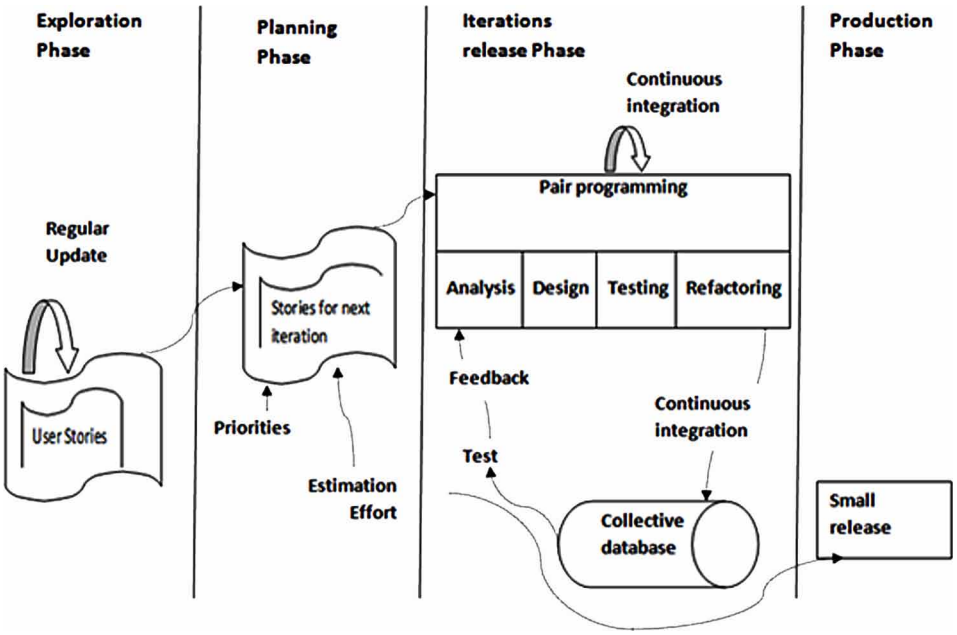
The core phases of the XP lifecycle as described by Beck (Beck, 1999) are discussed below- See Figure 2:

- **Exploration Phase:** In this phase, the requirements are collected. Customers write down the requirements that they intend to implement in the first release. The requirements are written on a card called a user story. Each user story card describes a set of features. In this phase, the project team identifies the tools, technologies and practices that will be used in the project. The duration of the exploration phase ranges from a few weeks to a few months, depending largely on how familiar the programmers are with the technology.
- **Planning Phase:** In this phase, the developers arrange the user stories cards by their priority level (i.e. which story is to be implemented first). Moreover, an agreement of the contents of the first small release is made. During this phase, an estimation is performed by programmers to estimate how much effort each user story requires and a schedule is then agreed upon. The duration of this phase is approximately two days.
- **Iterations to Release:** In this phase, several iterations of the system are conducted before the first small release. During iteration planning, the programmers break down the requirements or the features provided by the customers into a number of iterations that will each take one to four weeks to implement. At the end of the first iteration the architecture of the whole system is ready. This is achieved by selecting the stories that will enforce building the structure for the whole system. However, after the customer decides which stories are selected for each iteration, the programmers start, prior to coding the iteration, writing the test for the iteration. The programmers then create simple a design structure for the iteration. Acceptance tests are run by the customer at the end of every iteration. At the end of the last iteration, the system is ready for release.
- **Production / Release Phase:** In this phase, it is required from the development team to frequently make small releases of working software that customers can evaluate (Abrahamsson, Salo, & Ronkainen, 2002). This phase requires a lot of testing and checking of the performance of the system before the system can be released to the customer. During this phase, a small release of the system is given to the customer in order to test and check the functionality of the system and determine whether to accept the system or to provide feedback. The duration of this phase ranges from one week to three weeks.

The production/release phase is followed with the maintenance phase, which involves implementing the remaining requirements and updating the releases (including the new requirements of customer's feedback). After maintenance the final release is ready. This phase is known as the death phase.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Life cycle of the XP process



5. DESIGN METHODOLOGY

This section presents the design of an audit model that covers HIPAA security and privacy rules. The proposed auditing model is aligned with XP process. The design methodology in this section is based on an evaluation theory (Lopez, 2000): An Evaluation Theory Perspective of the Architecture Tradeoff Analysis Method (ATAM), and Common Criteria (Common criteria portal, 2012).

5.1. Preliminary Design Constraint

The Common Criteria for Information Technology Security Evaluation abbreviated as (Common Criteria or CC) have been selected to support the design of the proposed auditing model. CC is an international standard (ISO/IEC15408) that can be used to ensure and validate that a particular product or system meets a defined set of security requirements. The proposed auditing model, deploy specific standards from the CC that support HIPAA auditing rules which are illustrated below:

- Target of Evaluation (TOE) which can be done through the following standards:
  a. Protection Profile (PP)
  b. Security Target (ST)
- Security Function Requirements (SFRs)
- Security Assurance Requirements (SARs)
- Evaluation Assurance Level (EAL)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The requirements that were extracted from HIPAA privacy and security rules are: information system review, and audit controls. The information system reviews are related to the technical aspect of secure system development and the audit controls are related to the management aspect of secure system development. Therefore, mapping processes have been developed to select the CC standard that supports HIPAA information system reviews and audit controls. For instance, TOE could support the management and technical aspect of HIPAA auditing rules as shown in Figure 3.

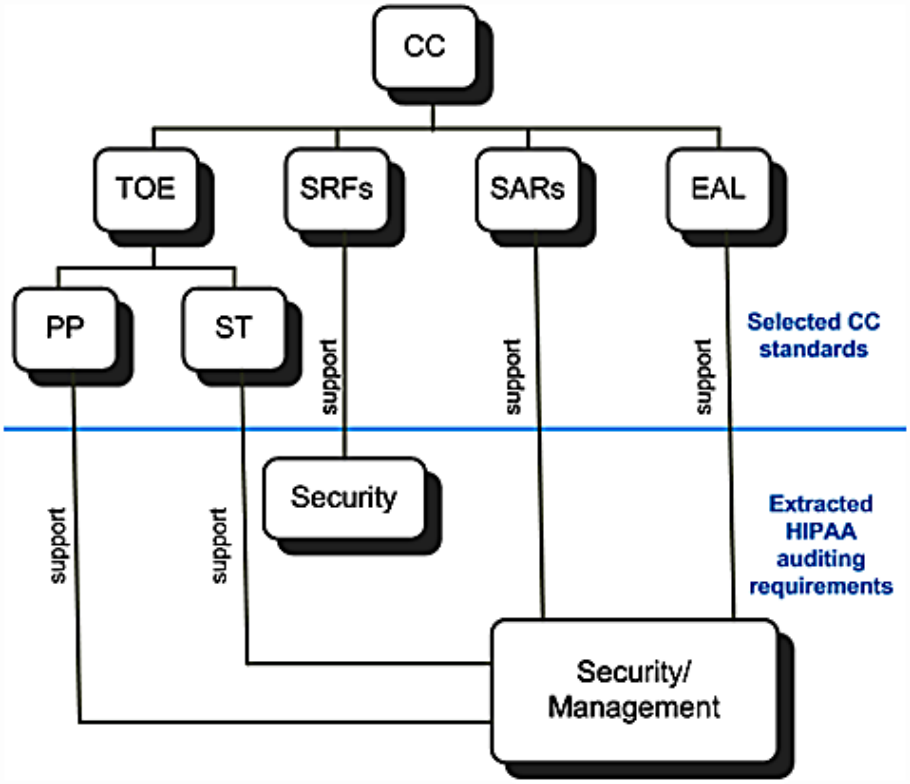
5.2. Evaluation Fundamental

The design of the proposed auditing model is based on the evaluation components proposed in (Lopez, 2000) and the evaluation components of CC as shown in Figure 4.

The main components of ATAM that will be used to support the design of HIPAA and the assessment of the auditing model are described as follows (Lopez, 2000):

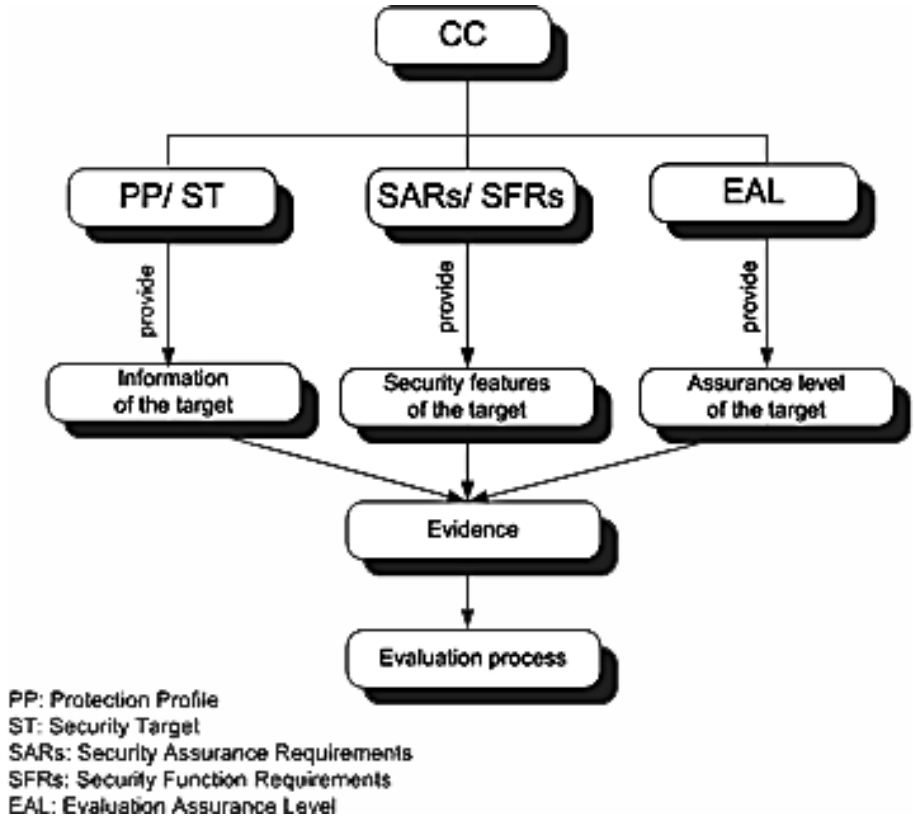
- **Target:** A system or product that is to be evaluated or audited.
- **Criteria:** The features of the target that is under evaluation or auditing.
- **Yardstick:** A comparison between ideal targets versus the actual targets based on specific criteria.
- **Data Gathering Techniques:** The process of gathering and measuring information for each criterion by using certain analysis techniques.

Figure 3. Mapping between CC and extracted HIPAA auditing requirements



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Evaluation components of CC



- **Synthesis Techniques:** The techniques used to prepare and synthesis information collected from analysis techniques, and then the results of synthesis compared to specific set of yardsticks.

The main components of CC that will be used to support the design of HIPAA and the assessment of the auditing model are described as follows:

- Selected standards of CC: PP, ST, SFRs, SARs, and EAL.
- **PP/ST:** Provide general information about the target to be audited. This information can be used as audit evidence.
- **SARs/SFRs:** Provide detailed information about the security features or security requirements such as security assurance requirements, and security functional requirements. This information can be used as audit evidence.
- **EAL:** Provide information about the assurance level of the evaluated target. This information can be used as audit evidence.

The design of the proposed auditing model follows a hybrid approach that combines the principles of Lopez- ATAM approach and the principles found in the CC, the first step of the design process is to identify the target of auditing and then identify the target's general information through PP/ST, and the target's security features through SFRs/SARs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

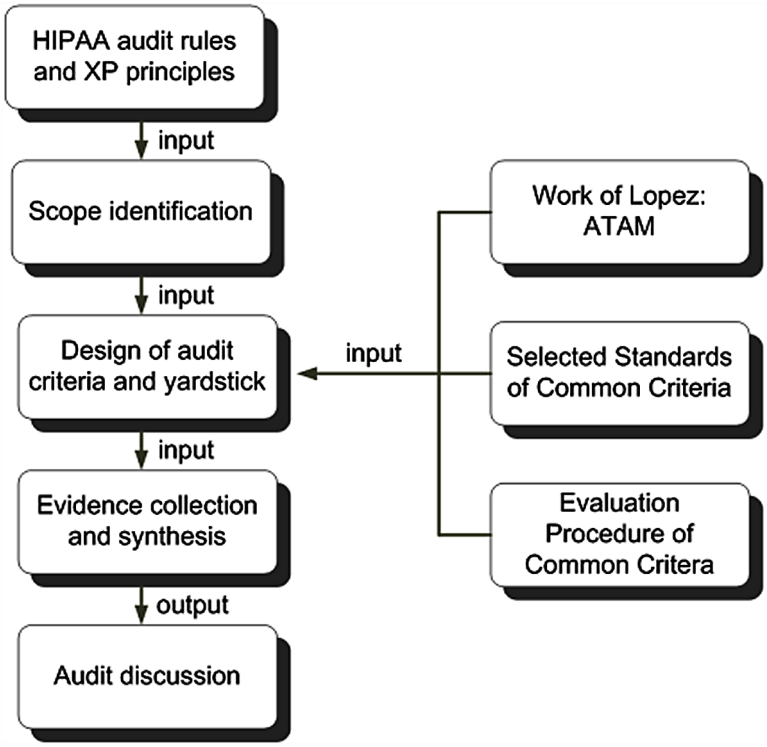
Lopez (2000) stated that when the target is already known and defined, its features must be identified for evaluation purposes. Once the data and security features are collected, the collected information is formulated in an applicable structure for the purpose of comparing it with the yardstick by using combination techniques, known as synthesis techniques. Figure 5 presents the main process of designing an audit model for HIPAA rules based on the work of Lopez and on the concepts of CC.

6. MODEL DESIGN

6.1. Scope Identification

The objective of the proposed auditing model for HIPAA security and privacy is to help HIPAA software auditors to extract the auditing evidences from the XP software processes. Additionally, the proposed audit model can also be useful for XP developers. It can help developers to follow the rules that are required by HIPAA and assure the auditors that certain steps and activities have been performed in compliance with HIPAA.

Figure 5. Design process for the audit model



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6.2. Design of the Audit Criteria and Yardsticks

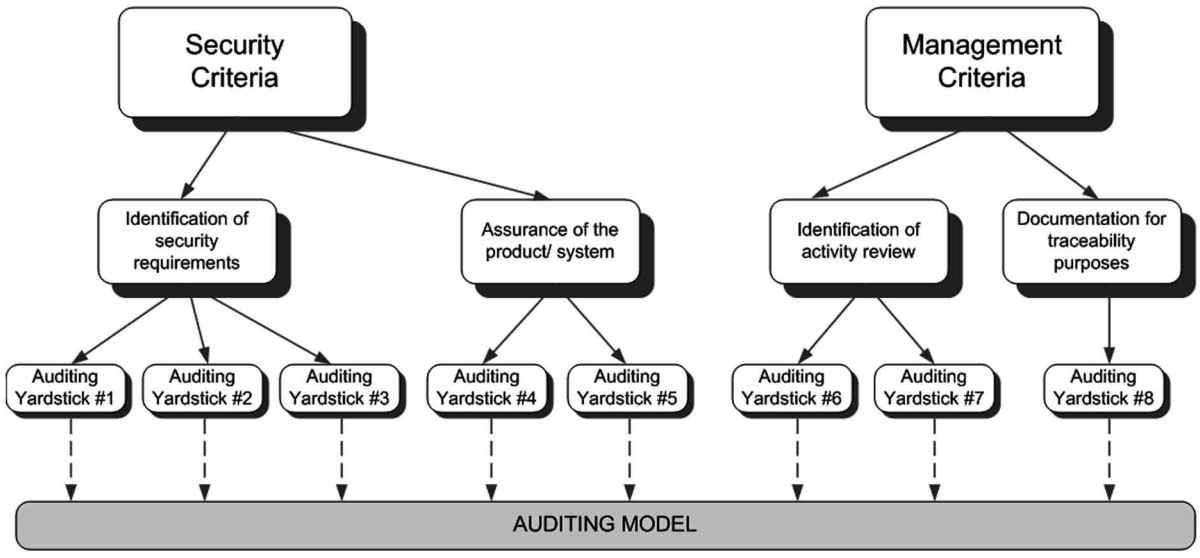
Auditing criteria is developed using a mandatory standard that indirectly contains the criteria to be applied in the auditing process. The auditing criteria should refer to any relevant study of targets, relevant standards, or ideals that might be relevant to the target in question (Qasaimeh, 2012; Lopez, 2000). Once the appropriate standards are selected, criteria should be driven from the selected standards. In this paper the mandatory standard is HIPAA. The proposed auditing model could not be achieved without the support from other relevant security standards such as ATAM and Common Criteria (CC) standard which identify appropriate processes for information technology security evaluation.

The objective of the auditing criteria is to translate the HIPAA security and privacy rules form legal based terminologies into a context based that can be understood for software project stakeholders. The proposed auditing model mainly consists of two categories of criteria: security criteria and management criteria and four main sub criteria (i.e. Auditing criterion) with multiple yardsticks for each criterion. The structure of the proposed auditing model is presented in Figure 6.

6.3. Security Criteria

The audit criteria presented in this section is based on the extracted HIPAA auditing rules and on the concepts of selected standards from CC such as PP, ST, SFRs, SARs, and EAL. The security criteria consist of two main auditing sub criteria which are identification of security requirements/features and assurance of the product/system

Figure 6. Structure of the proposed auditing model



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6.3.1. Identification of Security Requirements and Features

Identifying the security requirements is necessary and required by the HIPAA auditing rules. This can be achieved by using the concepts of selected standards of CC. The security requirements identified by the TOE represent the target to be evaluated which can be part of the system, a specific function of the system, or the whole system. Security features of the product or system are the core of HIPAA auditing rules. HIPAA focus on three security features: confidentiality, integrity, and availability (CIA). As stated in HIPAA rule: section 164.306 (a) (1) "Covered entities and business associates must ensure the confidentiality, integrity, and availability of all electronic protected health information the covered entity or business associate creates, receives, maintains, or transmits." The following are the audit yardsticks identified for this auditing criterion:

- **Yardstick #1:** This yardstick focuses on demonstrating that security objectives are documented by the project description. The focus on this yardstick is to extract the evidences typically created by consumer/user to represent the set of security objectives and features. The evidences for this yardstick can be obtained in several ways, one of which is the answer to the question of 'what I want or need' from intended secure system. This yardstick can be used also by the developer as a starting point for developing product from HIPAA viewpoints.
- **Yardstick #2:** This yardstick focuses on extracting evidences to identify that certain technique or methods have been used in the efforts of modeling the security requirements including detailed product specific information, security properties, and system specification.
- **Yardstick#3:** The objective of this yardstick is to extract security measures evidences from the software project. This yardstick focuses on the measures taken during development and evaluation of the product to ensure the compliance with the claimed security functionality. These assurance evidences involve design documentation, design analysis, functional testing, and vulnerability analysis.

6.3.2. Assurance of the Product /System

Assurance means the level of confidence for product that can be gained, to ensure that an entity will meet its security objectives, and to ensure that the security controls being obtained will operate correctly in the operational environment. The evaluation level of the product security is required by HIPAA. Therefore, the auditor should give an evaluation value to the system by assigning a number to the product or system. This number indicates the level of confidence of the system. This can be achieved by using the concept of Evaluation Assurance Level (EAL) found in CC standards. The following are the audit yardsticks identified for this auditing criterion:

- **Yardstick#4:** The objective of this yardstick is to extract the evidences related to the security function requirements from the project artifacts. This yardstick focuses on the software textual and design artifact such as investigating certain security functionality for an actual product or system, and the design decisions that specifies the services integration during the software development process. This yardstick mainly expresses 'what' a product shall do, and what shall not happen while executing the software.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Yardstick#5:** This yardstick focuses on extracting evidences that are belonging to the evaluation assurance level. Typically, a numerical grade is assigned to the product after the completion of the evaluation process. CC lists seven levels, from EAL1 to EAL7. The highest level provides the highest guarantee that the security features of the system or product are correctly applied. The EAL number assigned to the evaluated system indicate that the system implements all requirements successfully for that level.

6.4. Management Criteria

The management criteria that are required by HIPAA auditing rules are the reviewing activities and the documentation for project traceability, this requirement is mentioned in in section 164.308(a)(1)(ii)(c) "*implement procedures to regularly review records of information system activity*." The management criteria consist of two main auditing sub criteria which are identification of activity review method and documentation for traceability purpose.

6.4.1. Identification of Activity Review Method

Organizations that met or are aiming to meet the HIPAA standards, must have documented review process for the software development activities to provide evidence to the HIPAA's auditors that certain steps have been done correctly and comply with regulation requirements. The following are the audit yardsticks identified for this criterion:

- **Yardstick#6:** The objective of this yardstick is to extract the evidences to prove that the organization follows the customer requirements. This yardstick can be extracted by ensuring that the required services are implemented in the project.
- **Yardstick#7:** This yardstick focuses on evidences about the evaluation of potential security threats and describes the assets to those threats. This yardstick then investigates the existence of security controls and demonstrates that these controls are sufficient to counter the potential security threats.

6.4.2. Documentation for Traceability Purpose

Documentation and traceability is an important aspect for HIPAA auditing process that enables the software developers to find the origin of each requirement and track every change that was made to this requirement. For this purpose, it may be necessary to document every change made to the requirement.

An auditing story card is proposed to improve the agile stand-up meeting in order to capture the auditing evidences in a formal manner. The improvement to the stand-up meeting can be achieved by creating a WHAT, HOW, NEXT card as shown in Table 1, where the programmer writes down a short description of what he/she has done, how he/she has done it, and what is he/she going to develop next. The following are the audit yardsticks identified for this criterion:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. WHAT, HOW, NEXT card

Card elements	Description
WHAT	What has done first and which part of the software has been completed
HOW	The techniques to develop particular software part
NEXT	Which part of software will be integrated next

- **Yardstick#8:** The objective of this yardstick is to extract evidences that are related to the techniques improved by the developers to support the requirements traceability. This yardstick enables the auditor to find the origin of each requirement and track every change that was made to this requirement. The proposed user story card describes the type of user, which feature is needed to implement and why.

7. ASSESSMENT OF PROPOSED AUDITING MODEL

7.1. Context and Scope

According to ISO/IEC 14102 (2008), assessment in software engineering is an action of applying specific documented criteria to a specific software module, package or product for the purpose of software evaluation and enhancement. The purpose of assessment is to improve the quality of future software performances (Idomir, Idomir, & Repanovici, 2011). To assess the applicability of the proposed auditing model, four case studies have been selected as shown in Table 2. The selection of these case studies is based on the following steps and criteria:

Table 2. Selected case studies

Cases	Authors	Case title	Case objective	Case date
Case A	Kalra	Clinical Foundations and Information Architecture for the Implementation of a Federated Health Record Service	To propose a Federated Health Record (FHR) to be deployed in a live clinical environment in the Department of Cardiovascular Medicine at the Whittington Hospital in North London. However, implantation of Federated Health Record (FHR). The implementation of the proposed system is based on agile software development	2003
Case B	Sun, Zhu, Zhang, & Fang	HCPP: Cryptography Based Secure EHR System for Patient Privacy and Emergency Healthcare.	To propose secure EHR system and HCPP (Healthcare system for Patient Privacy), based on cryptographic methods and wireless network infrastructures	2011
Case C	Yepes, Cobaleda,Villa D., & Aedo	Design a Medical Application for Android Platform using Model-Driven Development Approach	This case proposed a method to design medical device interoperability system, called SIMMIT (System Integration Medical Monitoring and Interoperability for Tele-care) based on a Model-Driven Development approach	2012
Case D	Kalra, et al.	Security and Confidentiality Approach for the Clinical E-Science Framework (CLEF)	This case summarizes the overall approach adopted by CLEF to meet data protection requirements, that's required by HIPAA.	2003

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- A search was conducted on IEEE Xplore and ScienceDirect-Elsevier for any paper proposing a methodology, model, technique, or framework that improves a HIPAA compliant system or building medical based software system for specific use with agile software processes. The following terms were used in the search process:
  a. Agile and healthcare regulations
  b. Agile and HIPAA
  c. Agile-XP and critical projects and/or security projects
  d. XP and development of software medical device
  e. XP and implementation of electronic health record (HER)
- A title and abstract analysis was performed to select the papers that discuss XP and the implantation of healthcare projects such as software medical device and electronic health record servers.
- All the papers that discuss the implementation of critical projects or security projects that are not related to healthcare system were discarded.
- Some authors discuss the XP implementation or the design process of critical projects and security projects that use ePHI such as HER, these kinds of paper were selected.
- Some authors discuss the use of XP principals in developing software systems for medical devices and refer to these principles as agile processes without specifying which agile processes they are using.

7.2. Discussion of the Evidence Gathered

The international standard of auditing (2009) defines the audit evidence as "all the information used by the auditor in arriving at conclusions on which the audit opinion is based." In this research, the auditing evidences were classified based on the evidences found in (Qasaimeh, 2012) and which are summarized below:

- **Textual Evidence:** Can be described as supporting proof in certain text format.
- **Modeling Evidence:** Can be described as supporting proof in certain modeling format such as Unified Modeling Language (UML).
- **Graphical Evidence:** Can be described as supporting proof in certain graphical format such as bar graphs, line graphs, pie charts, scatter plots, photographs, and figures.

This section discusses the evidences that were gathered from the four selected case studies. Table 3 shows the summary of evidences found in the selected case studies. Below is the discussion of auditing evidences found on each case study.

Most of the evidences found in Case A demonstrate the security criteria of the proposed auditing model such as yardstick#2 which focuses on extracting detailed security requirements from the project. The auditing evidences for this yardstick are found in a modeling format (i.e. Use Case Diagrams). Case A also supports the management criteria of the proposed auditing model for example; yardstick#6 focuses on extracting the evidences to prove that the organization follows the customer requirements. The evidences for this yardstick are found in a graphical format, (e.g. screen shots of the system). Case A supports both the security criteria and the management criteria of the proposed auditing model.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Summary of evidences in the selected case studies

	Case A (Kalra, 2003)	Case B (Sun, et al, 2011)	Case C (Yepes, et al, 2012)	Case D (Kalra, et al, 2003)
Security criteria				
Yardstick#1 (Identification of security objectives)	Evidence Exists	Evidence Exists	Evidence does not exist	Evidence exists
Yardstick#2 (Detailed security requirements)	Evidence Exists	Evidence Exists	Evidence Exists	Evidence exists
Yardstick#3 (Security measures)	Evidence Exists	Evidence Exists	Evidence Exist	Evidence exists
Yardstick#4 (Services integration	Evidence Exists	Evidence Exists	Evidence does not Exists	Evidence exists
Yardstick#5 (Evaluation assurance procedures)	Evidence Exists	Evidence Exists	Evidence does not exist	Evidence exists
Management criteria				
Yardstick#6 (Organization reviewing activities)	Evidence Exists	Evidence Exists	Evidence does not exist	Evidence does not exist
Yardstick#7 (Identification of security threats)	Evidence Exists	Evidence does not exist	Evidence does not exist	Evidence exists
Yardstick#8 (Security requirements traceability)	Evidence Exists	Evidence exists	Evidence Exists	Evidence does not exists

Case B supports the security criteria of the proposed auditing model. As an example, it's written in Case B that "no one should be able to obtain the secrets for retrieving and decrypting the PHI" this statement could be considered as security/privacy rules that can be used as textual evidence to support yardstick#1. Case B partially support the management criteria of the proposed auditing model such as yardstick#8 which is focuses to extract evidences that are related to traceability of the project, this yardstick represented in case B as modeling evidence in UML format (i.e. Flowchart) that describe the construction of the secure system proposed by case B. However, no evidence can be found to support yardstick#6 and yardstick#7 of the management criteria for case B. As a result, case B supports the security criteria and partially supports the management criteria of the proposed auditing model.

The auditing evidences found in case C partially support both the security criteria and management criteria. Case C lacks the evidences of security criteria that required by proposed auditing model. For example, yardstick#1 does not exist which focuses on extracting evidence from the project to demonstrate that security objectives are in the project. Case C also lacks the evidences of management criteria that are required by the proposed auditing model such as yardstick#6, yardstick #7. Therefore, Case C does not comply with most of the proposed auditing yardsticks.

Evidences in Case D fully support the security criteria of the proposed auditing model. For example, all the yardsticks of the security criteria exist in Case D. But, the yardsticks of the management criteria of the proposed auditing model do not exist in case D. For that, Case D is rich in security aspects that are required by the proposed auditing mode and poor in management aspects.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

7.3. Detailed Auditing

The results shown in Table 4 indicate that the approach proposed by Case A could comply with HIPAA privacy and security requirements. The eight proposed yardsticks exist and are supported by Case A. For example, in regards to yardstick#1 which focuses on extracting evidence from the project to demonstrate that security objectives are in the project; this yardstick exists as textual evidence in section 6.2. (Structured set of requirements), and as graphical evidence in Table 7 page 208 of Case A. Another example is yardstick#3, which focuses on extracting evidences that are related to the security function requirements from the project, and also exists as textual evidence in Section 6.3. (FHR functional requirements) page 171 and Section 6.6. (FHR technical requirements) page 193 and as modeling evidence in Figures: 70, 71, 72, 73, 74, and 75 in the pages numbered 251 through to 253.

Table 4. Auditing results of Case A

Case A (Kalra, 2003)				
	Evidence Status	**Evidence Type**	**Evidence Location**	**Page number**
Yardstick#1 (Identification of security requirements)	Evidence exists	1-Textual evidence 2- Graphical evidence	1-Section 6.2. (structured set of requirements) 2- Table 7 (Examples of end users and their potential interactions with an FHR service)	1-P. 170 2- P.208
Yardstick#2 (Identification of security requirements)	Evidence exists	1-Modeling evidence 2-Graphical evidence	1- 6.10.1. (Use Case Diagrams) 2-Figures:38,39,40,41,42,43,44, 45	1-P. 203 2-P. 198 to 202
Yardstick#3 (Security features)	Evidence exists	1-Textual evidence 3-Modelling evidence	1-Section 6.3. (FHR functional requirements) 2-Section 6.6. (FHR technical requirements) 3-Figures: 70,71,72,73,74,75	1-P.171 2- P.193 3- P. 251 to 253
Yardstick#4 (Security features)	Evidence exists	1-Textual evidence 2-Textual evidence	1-Section 6.4.1 (Confidentiality and access control) 2-Section 6.4.5 (User Authentication) 3-Section 6.4.6 (FHR security)	1-P. 175 2-P.178 3-P.178
Yardstick#5 (Assurance of the product / system)	Evidence exists	1-Textual evidence	1-Section 6.4. (FHR medico-legal and security requirements)	1-P. 174
Yardstick#6 (Identification of activity review method)	Evidence exists	1-Graphical evidence 2-Textual evidence 3-Textual evidence	1-Figures 76 and 77 2-Section 11.3. (FHR Middleware Services) 3-Section 12.5. (Implementation and demonstration of the FHR service)	1-P. 257-258 2-P. 269 3-P.294
Yardstick#7 (Identification of activity review method)	Evidence exists	1-Textual evidence 2- Graphical evidence	1-Section 11.2. (Engineering approach for the record server middleware) 2-Figure 84	1-P.268 2-P.272
Yardstick#8 (Documentation for traceability purpose)	Evidence exists	1-Textual evidence 2-Modeling evidence	1-(The structure of health record entries STRC.9) 2- Figure 61	1-P. 185 2-P. 219

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As shown in Table 5, the EHR system and HCPP system proposed by Case B could comply with HIPAA security and privacy requirements. Seven out of eight yardsticks exist and are supported by Case B. For example, yardstick#2, which focuses on extracting detailed security requirements from project, exists as textual evidence in Section III and as graphical evidence in Figure 1 page 4. Another example is yardstick #3, which focuses on extracting security function requirements from the project to demonstrate that security objectives are described in the project's documentation also exists as textual evidences in Section I.B.1, page number 2, and Section III.C in page number 4.

Case C is a proposed method to design a system for medical device interoperability, called SIMMIT (System Integration Medical Monitoring and Interoperability for Tele-care) as shown in Table 6, Case C does not comply with HIPAA, because it lacks the security aspects required by HIPAA. For example, yardstick#3, which focuses on extracting evidence related to the security function requirements form the project, does not exist. Moreover, Case C lacks the management aspects required by HIPAA and which are represented by yardstick#6, yardstick #7.

Case D is rich in security aspects as required by HIPAA and poor in the managerial aspects. As shown in Table 7, most of the yardsticks that are related to the security aspects required by HIPAA exist, such as yardstick#2 that focuses on extracting evidence that contains detailed security requirements from the project and that can be considered as evidence of compliance. Yardstick#2 exists as graphical evidence in Figure 1. Management requirements that are required by HIPAA do not exist in case D. For example, yardstick#9 does not exist, which is provides evidences of traceability of the project.

Table 5. Auditing results of Case B

Case B (Sun, et al, 2011)				
	Evidence Status	**Evidence Type**	**Evidence Location**	**Page number**
Yardstick#1 (Identification of security requirements)	Evidence exists	1-Textual evidence	1-Section I.B	1-P.2
Yardstick#2 (Identification of security requirements)	Evidence exists	1-Textual evidence 2-Graphical evidence	1-Section III 2-Figure 1	1-P.3 2-P.4
Yardstick#3 (Security features)	Evidence exists	1-Textual evidence 2- Textual evidence	1-Section I.B.1 and 2 2-Section III.C	1-P.2 2-P.4
Yardstick#4 (Security features)	Evidence exists	1-Textual evidence	1-Section II.A.B and C	1-P.2
Yardstick#5 (Assurance of the product / system)	Evidence exists	1-Textual evidence	1-Section VI.A.B and C	1-P.9 and P.10
Yardstick#6 (Identification of activity review method)	Evidence exists	1-Textual evidence	1-Section V.A	1-P.7
Yardstick#7 (Identification of activity review method)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#8 (Documentation for traceability purpose)	Evidence exists	1-Textual evidence 2-Modeling evidence	1-Section IV.A 2-Figure 2	1-P.5 2-P.5

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Auditing results of Case C

Case C (Yepes, et al, 2012)				
	Evidence Status	**Evidence Type**	**Evidence Location**	**Page number**
Yardstick#1 (Identification of security requirements)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#2 (Identification of security requirements)	Evidence exists	Modelling evidence	Figure 2 Figure 4 Figure 6	P.3 P.4 P.4
Yardstick#3 (Security features)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#4 (Security features)	Evidence exists	Modelling evidence	Figure 3 Figure 5	P.4 P.4
Yardstick#5 (Assurance of the product / system)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#6 (Identification of activity review method)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#7 (Identification of activity review method)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#8 (Documentation for traceability purpose)	Evidence exists	Textual evidence	Section 4	P.2

Table 7. Auditing results of Case D

Case D (Kalra, et al, 2003)				
	Evidence Status	**Evidence Type**	**Evidence Location**	**Page number**
Yardstick#1 (Identification of security requirements)	Evidence exists	Textual evidence	Section 3	P.2
Yardstick#2 (Identification of security requirements)	Evidence exists	Graphical evidence	Figure 1	P.2
Yardstick#3 (Security features)	Evidence exists	Graphical evidence	Figure 2	P.3
Yardstick#4 (Security features)	Evidence exists	Textual evidence	Section 4.3	P.5
Yardstick#5 (Assurance of the product / system)	Evidence exists	Textual evidence	Section 4.4	P.5
Yardstick#6 (Identification of activity review method)	Evidence does not exist	*Not applicable*	*Not applicable*	*Not applicable*
Yardstick#7 (Identification of activity review method)	Evidence exists	Textual evidence	Section 5	P.7
Yardstick#8 (Documentation for traceability purpose)	Evidence does not exists	*Not applicable*	*Not applicable*	*Not applicable*

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Impacts of the proposed auditing model on agile XP phases

Agile XP phase	Impacted By
Exploration phase	Yardstick#1 (security objectives) Yardstick#2 (detailed security requirements) Yardstick#3 (security function requirements)
Planning phase	Yardstick#4 (security assurance requirements) Yardstick#5 (evaluation assurance level) Yardstick#7 (risk analysis)
Iterations to release phase	Yardstick#6 (meets customer requirements) Yardstick#8 (traceability)
Production / Release Phase	Yardstick#5 (evaluation assurance level) Yardstick#6 (meets customer requirements)

8. IMPACTS OF PROPOSED AUDITING MODEL ON AGILE XP

The core phases of XP life cycle are: exploration, planning, iterations to release, and production/ release phase. As shown in Table 8 and discussed below, the proposed auditing model will have an impact, either direct or indirect, on each phase of agile XP process.

- **Exploration Phase:** This phase focuses on collecting requirements from the customer. If the XP developer follows *Yardstick #1 (security objectives)* which focuses on extracting the evidences typically created by the consumer/user to represent the set of security objectives and requirements needed, the developer will identify the requirements needed for HIPAA. Moreover, by following *Yardstick #2 (detailed security requirements)* and*Yardstick#3 (security function requirements),* the developer will be able to recognise the details of the security requirements that need to be submitted as HIPAA compliance evidence. The detailed security requirements include specific product details such as security properties, and product or system specification.
- **Planning Phase:** During this phase, time, cost and risk estimation is performed by the programmers. Specific criteria, such as *Yardstick#4* which focuses on the measures taken during development and evaluation of the product to ensure compliance with the claimed security functionality, *Yardstick#5* which aims *to evaluate the assurance level* and *Yardstick#7* which provides criteria for *risk analysis,* all have an impact during this phase.
- **Iterations to Release Phase:** In this phase, several iterations of the system are conducted before the first small release. Acceptance tests are run by the customer at the end of each iteration. At the end of the last iteration, the system is ready for release. Similarly, several criteria have an impact on this phase including *Yardstick#6 (meets customer requirements)* which focuses on extracting evidences to prove that the organization follows the customer requirements and *Yardstick#8 (traceability)* which focuses on extracting evidences that are related to traceability of the project.
- **Production / Release Phase:** During this phase, more testing and evaluation are conducted before the system is released to the customer. During this phase, a small release of the system is given to the customer to test and check the functionality of the system in order to determine whether to accept the system or to provide feedback. Both *Yardstick#6 (meets customer requirements)* and *Yardstick#5 (evaluation assurance level)* which tests how secure the product is, have an impact on this phase.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

9. CONCLUSION AND FUTURE WORK

The aim of this paper is to enhance the auditability of XP in response to HIPAA rules, as well as to help XP based software organizations that develop healthcare software in their effort to become HIPAA certified. This paper proposes an auditing model for HIPAA security and privacy rules in XP environments. The design of the model is based on an evaluation theory which takes as its input the work of Lopez (ATAM), and standards of Common Criteria (CC) concepts, as well as an evaluation procedure proposed by CC concepts. Finally, the proposed auditing model has been assessed based on four case studies to investigate the compliably of the selected case studies with the HIPAA security and privacy rules.

The main limitation of this research is the assessment approach which has been carried out by the same researchers who designed the model. This might introduce biases in the assessment process of the proposed model. For future work, the validation of the model is planned to be performed by external experts to assess the usability of the proposed model in the environment of XP. In this way, the same case studies can be audited by bodies external from the research team. This research as can also be improved by conducting the proposed audit model on medical projects that already implemented by agile software organization and wishing to become HIPAA compliant

REFERENCES

AAMI. (2012). Guidance on the use of agile practices in the development of medical device software. *AAMI TIR45*. Retrieved from http://www.aami.org

Abrahamsson, P., Salo, O., & Ronkainen, J. (2002). Agile software development methods: Review and analysis. Espoo VTT Publications, 43(9), 10-12.

Ahmed, S. (2014). Simulation method to improve hospital service quality. *International Journal of Information Systems in the Service Sector*, *6*(3), 96–117. doi:10.4018/ijisss.2014070106

Beck, K. (1999). Embracing change with extreme programming. *IEEE Computer*, *32*(10), 70–77. doi:10.1109/2.796139

Brian, K., & Daniel, M. (1997). The politics of the health insurance portability and accountability act. *Health Affairs*, *16*(3), 146–150. doi:10.1377/hlthaff.16.3.146 PMID:9141331

Common Criteria Portal. (2012). Common criteria for information technology security evaluation, part 1: Introduction and general model. Retrieved from https://www.commoncriteriaportal.org/files/ccfiles/CCPART1V3.1R4.pdf

Fitzgerald, B., Stol, K., O'Sullivan, R., & O'Brien, D. (2013). Scaling agile methods to regulated environments: An industry case study. *Proceedings of the 35th International Conference on Software engineering*, San Francisco, CA, USA. 10.1109/ICSE.2013.6606635

Idomir, M., Idomir, V., & Repanovici, F, A. (2011). Web-based application for self-testing and assessment. *International Journal of Education and Information Technologies*, *5*, 141–148.

International Federation of Accountants. (2009). International standard on auditing 500. Retrieved from http://www.ifac.org/sites/default/files/downloads/a022-2010-iaasb-handbook-isa-500.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ISO IEC 14102. Information technology- guideline for the evaluation and selection of case tools. (2008). International Organization for Standardization. Geneva, Switzerland.

Kalra, D. (2003). Clinical foundations and information architecture for the implementation of a federated health record service [Doctoral thesis]. University of London.

Kalra, D., Singleton, P., Ingram, D., Milan, J., MacKay, J., Detmer, D., & Rector, A. (2003). Security and confidentiality approach for the clinical e-science framework (CLEF). *Proceedings of the Second UK E-Science All Hands Meeting*, Nottingham, UK (pp. 825-832).

Lopez, M. (2000). An evaluation theory perspective of the architecture tradeoff analysis method (Technical Report CMU/SEI-2000-TR-012). Software Engineering Institute, Carnegie Mellon University.

McHugh, M., Mc Caffery, F., & Casey, V. (2014). Adopting agile practices when developing software for use in the medical domain. *Journal of Software: Evolution and Process*, *26*(5), 504–512.

McHugh, M., Mc Caffery, F., Fitzgerald, B., Stol, K., Casey, V., & Coady, B. (2013). Balancing agility and discipline in a medical device software organization. *International SPICE Conference Process Improvement and Capability determination*, Germany. 10.1007/978-3-642-38833-0_18

McHugh, M., McCaffery, F., Casey, V., & Pikkarainen, M. (2012). Integrating agile practices with a medical device SDLC. *Proceedings of the European Systems and Software Process Improvement and Innovation Conference* EuroSPI, Vienna, Austria.

Mehrfard, H., Pirzadeh, H., & Hamou-Lhadj, A. (2010). Investigating the capability of agile processes to support life-science regulations: The case of xp and fda regulations with a focus on human factor requirements. *Proceedings of the 8th ACIS International Conference on Software Engineering Research, Management and Applications*, Montreal, Canada.

Portability, H. I., & Act, A. (1996). *HIPAA* (pp. 104–191).

Qasaimeh, M. (2012). *Auditing for ISO 9001 requirements in the context of agile software processes.* [Unpublished doctoral dissertation]. University of Quebec, Montreal, Canada.

Qasaimeh, M., & Abran, A. (2011). Extending extreme programming user stories to meet ISO 9001 formality requirements. *Journal of Software Engineering and Applications*, *4*(11), 626–638. doi:10.4236/jsea.2011.411074

Qasaimeh, M., & Abran, A. (2013). An audit model for ISO 9001 traceability requirements in agile-xp environments. *Journal of Software*, *8*(7), 1556–1567. doi:10.4304/jsw.8.7.1556-1567

Reina, R., Lacroce, A., Cetani, T., & Ventura, M. (2012). A healthcare project: Managing knowledge through electronic medical record-empirical cases. *International Journal of Information Systems in the Service Sector*, *4*(4), 1–19. doi:10.4018/jisss.2012100101

Sun, J., Zhu, X., Zhang, C., & Fang, Y. (2011). HCPP: Cryptography based secure EHR system for patient privacy and emergency healthcare. *Proceedings of the IEEE International Conference on Distributed Computing Systems* (pp. 373–382). 10.1109/ICDCS.2011.83

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

U.S. Department of Health & Human Services. (2014b). Health information privacy. business associates. Retrieved from http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/businessassociates.html

U.S. Department of Health and Human Services. (2014a). Health information privacy. for covered entities and business associates. Retrieved from http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/

Wall, W. (2009). Developing global competitiveness in healthcare: A Thai healthcare organization's perspective. *International Journal of Information Systems in the Service Sector*, *1*(4), 61–72. doi:10.4018/jisss.2009062905

Westfall, L. (2007). Software configuration management audits. *The WestFall Team.* Retrieved from http://www.westfallteam.com/node/52

Wiegers, K. (2002). *Peer reviews in software: A practical guide*. Boston, MA: Addison-Wesley.

Yepes, J., Cobaleda, L., & Villa, D., & Aedo. (2012, July 16-19). A model-driven development approach to design software for embedded systems using the android platform. *Proceedings of the International Conference On Software Engineering Research & Practice*, Las Vegas, USA (pp. 45-52).

This research was previously published in the International Journal of Information Systems in the Service Sector (IJISSS), 9(1); edited by John Wang; pages 1-21; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 68
The Cloud Gets Personal:
Perspectives on Cloud Computing for Personalized Medicine

Rhoda Joseph
Penn State University, USA

Patrick Brown
Thermo Scientific NanoDrop Products, USA

ABSTRACT

Cloud computing, at its most fundamental level, is a technology that allows computing to occur as an on-demand utility. Personalized medicine is revolutionizing the field of medicine, and aims to provide individualized treatment for patients, based on their genetic material and medical history, through comparisons to genetic material from thousands of other patients. Better patient outcomes are expected in a personalized medicine domain as compared to more generic generalized options. In this paper the authors explore the use of cloud computing in the area of personalized medicine examining unique benefits and opportunities. They also recognize and discuss the presence of some inherent risks. Using the cloud for personalized medicine can expedite collaborations and data collection across different clinicians, researchers and other stakeholders involved in the fight against diseases. This paper examines key elements associated with using cloud computing in the specialized area of personalized medicine.

1. INTRODUCTION

The top three causes of death in the United States are heart disease, cancer, and chronic lower respiratory diseases (CDC, 2013). Each year thousands of people suffer through and die from chronic diseases. Information and communications technologies (ICTs) have long been a stalwart in the healthcare industry for diagnosing, treating and preventing chronic diseases. Wearable devices, sensors, mobile technologies, electronic health records (EHRs) are a few of the technologies impacting healthcare today (Jayanthi, 2014). In this paper, we demonstrate the use of cloud computing for personalized medicine as another tool in the arsenal of technologies to fight disease and improve human health.

DOI: 10.4018/978-1-7998-1204-3.ch068

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Two main research questions are examined in this paper: 1. what are the risks associated with the use of cloud computing in the area of personalized medicine; and 2. how does/can cloud computing impact the field of personalized medicine? Personalized medicine first requires a genetic profile of the patient. This profile provides very specific information about the patient and also the disease(s) that they are diagnosed with. This profile is then analyzed, and compared to thousands of profiles of other patients to find an optimal tailored solution for the current patient.

One recent study by Horowitz (2014) identified the value of cloud computing in the area of personalized drug treatments for patients. More specifically this study collected genetic data and examined the treatment of elderly patients that take a large combination of prescription drugs. The program in the study - PACE (Program of All-inclusive Care for the Elderly) is a Medicare and Medicaid program that partnered with two firms: Coriell Life Sciences and CareKinesis; used cloud computing along with personalized medicine to better serve elderly patients taking as many as 16 different prescription drugs per day (Horowitz, 2014). Anticipated outcomes of this program include better understanding of drug side-effects and faster responses to patients that may develop complications associated with drug cocktails.

The use of cloud computing for personalized medicine requires effective management of sensitive and vast datasets. Vast datasets are characteristic of the big data landscape that exists today. As data sets continue to grow, it is increasingly challenging to maintain privacy for genomic data that is shared across computer networks (Chow-White, MacAulay, Charters, & Chow, 2015). In addition to managing the challenges associated with large data sets other issues exist such as: managing the versatility across different environments; scaling to accommodate large data centers; automating to create self-managed environments; and integration of non-cloud products into the cloud architecture (Martin-Flatin, 2014). Personalized medicine involves collaboration across different entities, so overcoming some of the managerial challenges is paramount for continued growth and success.

Personalized treatments for patient conditions are already established in the personalized device industry. Currently, personalized medical devices using both iOS and Android technologies comprise a financially lucrative industry of approximately $27 billion (Bendor-Samuel, 2014). Examples of personalized medical devices include the “Tinnitus Masker” a product that is designed for an individual’s auditory needs; and the “Zenith Fenestrated AAA Endovascular Graft” a product that is customized to fit the patient’s aortic anatomy (U.S. Food and Drug Administration FDA, 2013). Undoubtedly there is a strong financial incentive for some related businesses to pursue personalized medicine. Hopefully, these financial business gains will also translate to improving the quality of patient lives and reducing premature mortality.

To develop the presented research questions, this paper proceeds in the following sections. First we present an overview of cloud computing. Then an overview of personalized medicine is presented followed by a discussion on the overlap of personalized medicine and cloud computing. The theoretical model is presented showing the key factors involved in examining their intersection. Lastly, discussion, implications and conclusions are presented.

2. CLOUD COMPUTING OVERVIEW

Cloud technology has become an important change agent in the field of information technology. Cloud computing has revolutionized the way that many businesses operate today, by freeing up internal resources, and shifting some technology related tasks and data management to external cloud computing

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

vendors. The existing market for cloud computing in the healthcare field is segmented into two main categories: clinical information systems (CIS) with examples such as electronic health records and bioinformatics; and non-clinical information systems (NCIS) that includes items such as billing and payroll (Newswire, 2015).

The term "cloud" has become part of the daily vernacular in the field of information systems. In its simplest form the cloud utilizes the power of the Internet to shift computing architecture including hardware and software from on-site machines to off-site computers. At a consumer level, an individual may store personal data such as documents, images and music on Google Drive or Apple iCloud. Many companies, both large and small, are contemplating the use of or are currently using cloud services; however, as with any advancement in technology, there are careful cost-benefit considerations that must be examined before implementation.

A business or organization must also consider which category of cloud computing is most appropriate for their needs. One method used to classify cloud computing is based on the specific type of service that is provided. The three main service level categories are: IaaS (Infrastructure as a Service) which occurs when a cloud vendor provides virtual/remote machines for clients to use; PaaS (Platform as a Service) which occurs when the cloud vendor provides a developmental space for the client to use; SaaS (Software as a Service) which occurs when the client remotely leases software from the vendor (Chang, Walters, & Wills, 2013; D. Wang, 2015). . Other terms in the cloud computing vernacular include: HaaS (hardware as a service) and DaaS (data as a service) both subcategories of IaaS (H. Yang & Tate, 2012).

Cloud computing environments are supported by a variety of different technologies ranging across multiple types of hardware, software, algorithms and models. Data storage in the cloud has increased the size and scope of the amount of data that individuals and organizations can collect. As part of the changing landscape of cloud computing hardware, Amazon has introduced two new products the Snowball Edge and Snowmobile (Forrest, 2016). Both products function as mobile "semi-trucks" that allows Amazon to physically migrate data to their AWS cloud solution. These products also have embedded relevant software that includes capabilities for large scale applications such as Internet of Things (IoT) and big data applications. Changes in cloud technological solutions will impact application areas such as personalized medicine.

At a minimum a cloud architecture consists of a cloud service provider and a cloud consumer, but can be extended to include a cloud broker, cloud developer or cloud auditor (Samba, 2012). The cloud broker functions as an intermediary between the provider and consumer; the cloud developer designs and implements services based on customer requirements; and the auditor independently assesses performance and operations of the cloud provider in the interest of the clients. Each of the above configurations can occur with a private cloud (an internal organization solution), public cloud (independent vendor(s) with cloud solutions available to the general public including organizations and individuals), community cloud (shared solution for a community) and hybrid cloud (combination of two or more from private, public or community cloud). In the medical field cloud computing has an existing established presence (Badger, Grance, Patt-Corner, & Voas, 2012; Newswire, 2015; Puustjärvi & Puustjärvi, 2013). In this paper we specifically examine the area of cloud computing as it pertains to personalized medicine.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3. PERSONALIZED MEDICINE

Personalized medicine, also called precision medicine, genomic medicine, information-based medicine, and individualized medicine; encompasses individualized, customized treatments based on patient genetic material and environmental variables. Researchers have identified hundreds of specific genes and mutations associated with different illnesses, such as specific genetic abnormalities associated with breast cancer or colorectal cancer (Hamburg & Collins, 2010). One of the earliest genetic testing for disease was for a genetic mutation causing Huntington's disease (a neurodegenerative disease) in the early 1990's; since then steady interest in personalized medicine has grown, and it is expected to continue to grow with increasing healthcare costs and an aging population (Bolouri, 2010).

Personalized medicine requires access to detailed patient information including patient medical history and DNA information. With the genome sequencing of thousands of individuals, sub groups can be used to compare a specific individual's data and screen for different illnesses; following this, clinicians can provide better predictions and outcomes regarding individual patient care (Bolouri, 2010). Some third party private companies market services directly to consumers who desire individual genetic testing (e.g. 23andMe, Navigenics and i-gene). Private companies provide information to their clients such as genetic ancestry and genetic traits.

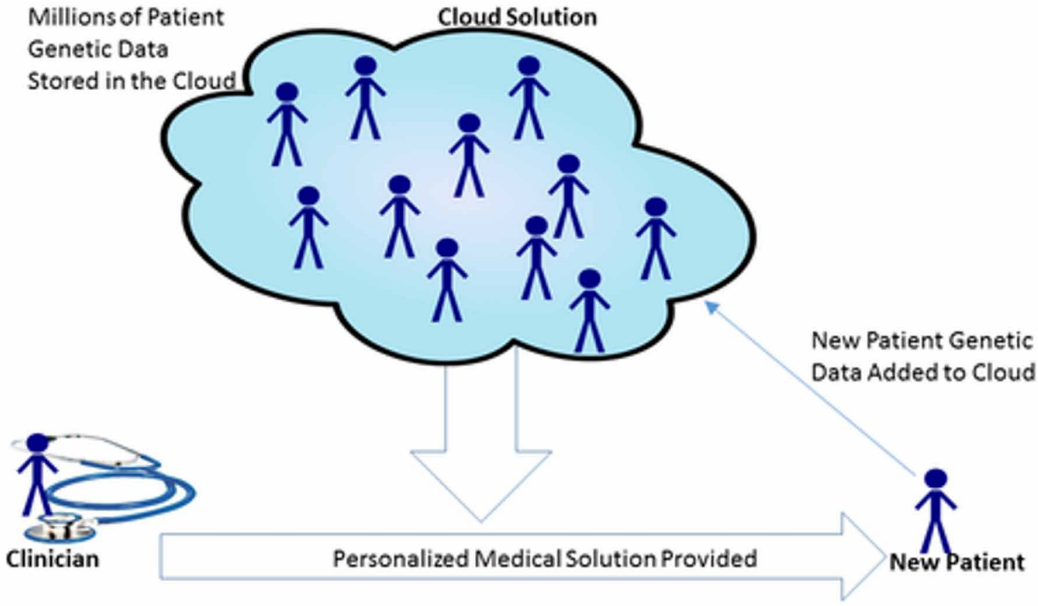
In this paper instead of looking at individual companies that are either marginally or directly involved in personalized medicine, we take a macro-level approach and look at the multiple partners involved in personalized medicine. The macro level includes both private and public entities such as patients, labs, hospitals, universities, government agencies and other agencies involved in personalized medicine projects. Even though there is tremendous potential for improved healthcare delivery through personalized medicine, its value has not been successfully communicated to relevant entities (Johnson, 2016). Effective collaborations, facilitated through relevant ICTs, can advance the case for personalized medicine.

Current medical treatments are prescribed based on how the average person responds to the dosage of medicine. Because of the variations between individuals, the same treatment dosage may work better or worse on different patients and yield different side effects. One of the key aspects of personalized medicine is the idea that treatment for a condition can be individualized or grouped with other similar patients, based on the patient's genetic traits compared to existing genomic datasets. As shown in Figure 1, the genetic material of each new patient is added to millions of records of other patients. Based on complex data computations combining the new patient data with the existing data, a personal medical profile for the new patient is created. The clinician can then use the personalized profile to find a more optimal solution to treat the new patient. Since the human genome was mapped in the early 2000s, interest in personalized medicine has grown tremendously and it is considered as a new cutting edge approach for medical treatment (Deakin, 2011; Podshivalov, Fischer, & Bar-Yoseph, 2014). However, large amounts of data are required for effecting personalized treatments.

The first step in personalized medicine is to gather and store patient data. When a new patient enters a hospital or medical facility, the patient must first consent to having his/her DNA purified from blood or saliva samples and having his/her genome mapped. A matrix of data is constructed including diagnosis, drug regiment, and genetic traits. The results of these analyses are stored in bio-repositories where researchers and clinicians can access them. Patient identifiers are removed in order to maintain patient confidentiality. The goal is that over time the drug treatment and genetic records of millions of patients

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Simplified clinician-patient personalized medicine cloud view



will be accessible to clinicians who can then compare their current patient to the database and determine the best course of treatment for the patient. Obstacles that need to be overcome include collection and storage of the data, accessibility of the data and querying of the data, and cost of testing.

As clinicians and other medical professionals add more individual patient information to relevant databases, more precise outcomes can be derived to treat individual patients. An individual's genetic information can range from family traits such as eye color, to more medically relevant information such as future risks for many diseases including different types of cancers. Personalized medicine is characterized as covering the 4Ps of medicine: personalized, predictive, proactive and participatory (Deakin, 2011).

Expansion of the 4Ps begins with personalization, which is the individualized care goal for each patient. Secondly, the computational power over thousands or even millions of collected patients' genetic profiles increases the predictive validity of future results. A proactive, solution oriented approach increases the accuracy of predicting the best course of action to treat individual patients Lastly, a key driver for success in the arena of personalized medicine is participation. If tools are available for easy upload of more data points, more gains can occur. Due to the network effect, in a clinical environment, as more patient information is entered, the value of the entire network increases.

The push towards personalized medicine requires that a clinician has access to a database that contains patient information either at a regional server or on a national server level. From there the clinician needs to query the database to determine the best treatment dosage for his or her patient given the prior patients' outcomes stored in the database. As the amount of information held in bio-repositories increases, the processing capabilities required to store and access the data will also increase beyond the capabilities of many existing stand-alone database systems. To solve this problem, partnerships between personalized medicine practitioners and cloud computing providers presents a viable solution.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. THE CLOUD FOR PERSONALIZED MEDICINE – THEORETICAL JUSTIFICATION

Successfully implementation of personalized medicine using cloud computing requires a committed collaborative approach from different participants including clinicians, labs, medical facilities, hospitals, research facilities and of course patients. Resource pooling, on-demand services and elasticity are core characteristics of cloud computing (Badger et al., 2012) that bode well for collaborations and flexibility involved in personalized medicine. Collaboration theory provides the theoretical lens to support the use of cloud computing for personalized medicine.

Collaboration theory (Gray, 1989) identifies three phases in a collaborative endeavor: 1. problem setting – define the problem and get the different stakeholders involved; 2. direction setting – create an agenda and agreement; 3. implementation – utilization of solution, building support and ensuring compliance by participants. For cloud computing and personalized medicine there are multiple actors involved including cloud provider(s), clinicians, hospitals, and other relevant entities. The second phase of collaboration theory requires the creation of an agenda and agreement amongst relevant parties. In the personalized medicine cloud domain there can be multiple agendas and agreements. Lastly, the implementation of the solution is an on-going endeavor.

To demonstrate collaboration theory in action consider the area of pharmacogenetics - the study of understanding the genetic profile of an individual's specific disease. First the problem setting: understanding the genetic profile of a patient's specific illness. Second, is the direction setting – creating an agenda for the use of personalized medicine in the area of pharmacogenetics. Pharmacogenetics is already poised to be one of the primary beneficiaries of personalized medicine (Deakin, 2011; Sheng, Li, & Wong, 2015; U.S. Food and Drug Administration FDA, 2013). Lastly, the third piece of collaboration theory deals with implementation of a solution. This third piece would require buy-in and participation from different entities such as patients, clinicians, medical facilities etc.

With personalized medicine the one size fits all for medical treatment approach is no longer the only option for patient care. Ideally, each patient will receive a customized and personalized analysis of their disease and a unique treatment for his/her illness. However, only through the inputting of data points from thousands of patients, and effectively processing that data will these outcomes be possible. Cloud computing provides an opportunity for disparate and geographically dispersed participants to share and integrate such data.

The presence of disparate silos of human genetic data in the medical field, presents a compelling argument in support of integrated cloud-based solutions for personalized medicine. A recent study identified and highlighted the varied sources of data that is collected for personalized healthcare including data from labs, clinicians, simulations, wearable sensors, and medical imaging (Viceconti, Hunter, & Hose, 2015). Table 1 provides a sample of large scale genomic data repositories. The existence of data silos limits sharing of critical genetic variants that can lead to better outcomes for patients, providers, and other relevant parties involved (Rehm et al., 2015). Using cloud computing for personalized medicine facilitates more data integration however there are implications (developed further in sections 5 & 6) for many stakeholders including clinicians, researchers, patients, medical facilities, insurance companies, pharmaceutical companies and others.

From a solution oriented perspective, both private and public entities can participate in the cloud strategy for personalized medicine. Further, private, public, community and hybrid cloud options make up the solution. A medical facility can use their internal data and integrate patient information through

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Selected sample of large scale genomic datasets

Name	Data	Overview	Website	Reference
The Cancer Cell Line Encyclopedia (CCLE)	Data for approximately 100 cell lines.	Data used for human cancer models	http://www.broadinstitute.org/ccle/home	(Barretina et al., 2012)
The National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS)	Contains genomic Alzheimer's datasets.	Portal with software, resources and genomic data about Alzheimer's.	https://www.niagads.org/	(U.S. National Library of Medicine, 2015)
National Database for Autism Research (NDAR)	Genomic data collected from over 100,000 subjects.	Data repository for Autism research such as behavioral, genetic, environmental, neural tissue data.	https://ndar.nih.gov/	(U.S. National Library of Medicine, 2015)
Genomics of Drug Sensitivity in Cancer	Data for over 1000 human cancer cell lines.	Data used for human cancer models	http://www.cancerrxgene.org/about/	(Yang et al., 2013)
Global Alzheimer's Association Interactive Network (GAAIN)	Data from global partners including genetic information.	Shared network studying the causes of Alzheimer's disease.	http://gaain.org/	(Neu, Crawford, & Toga, 2016)

partnerships with other institutions. For example the laboratory for personalized medicine (LPM) established in 2009 at Harvard Medical School, uses Amazon's (AWS) cloud computing architecture to develop individual solutions for patients through computer simulations of millions of avatars (Center for Biomedical Informatics (CBMI) at Harvard Medical School, 2014). Both public and private sector entities are key participants in the continued upward trajectory for the growth of personalized medicine.

5. CHALLENGES OF CLOUD COMPUTING FOR PERSONALIZED MEDICINE

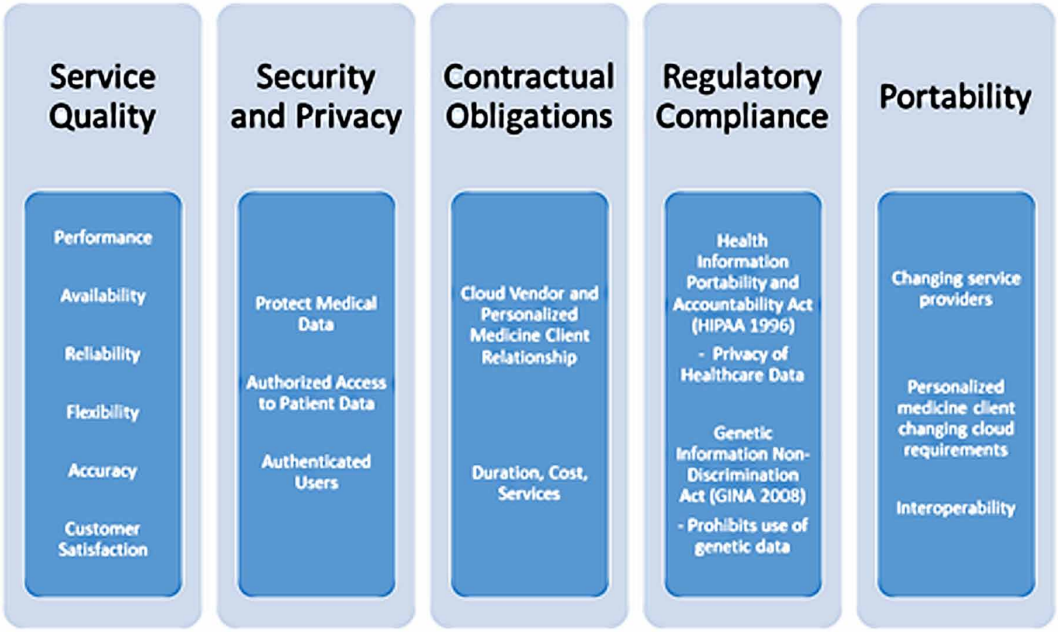
Although cloud computing has many benefits, there are also risks to using off-site technologies. Additionally, even though personalized medicine is an exciting area for innovation, there are some concerns as this field grows. The main challenges we identified associated with the use of cloud computing for personalized medicine are: service quality; security and privacy concerns; contractual obligations; government regulations; and portability. Each of these factors is described in more detail in Figure 2.

5.1. Service Quality

Unlike in-house software and hardware products where the business controls all its IT, with cloud-computing the vendors are in control of some or all of the IT infrastructure. As a result of this relationship dynamic, the quality of service provided by the vendor is an important issue. Three key dimensions of cloud service quality are performance, availability and reliability (Ardagna, Casale, Ciavotta, Pérez, & Wang, 2014). Particularly in the personalized medicine domain the system must perform in a manner consistent with the defined expectations. The systems must also be available whenever needed.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Risks of cloud computing for personalized medicine



In addition to the above three components of quality, we added three more variables: flexibility, accuracy and customer service. Flexibility reiterates the on-demand nature of cloud computing. Accuracy whether in SaaS, PaaS, or IaaS is paramount. Lastly, customer feedback and interaction can he used to improve the quality of service. Effective mechanisms to capture and resolve user complaints will also add to the perceived value of the cloud services provided.

5.2. Security and Privacy

The second risk with the cloud for personalized medicine is that of security and privacy. The cloud vendor must meet security and privacy standards for all sensitive data, even in the presence of anonymized patient data. Security and privacy ought to be of utmost concern to a client contemplating using cloud-computing technology in the area of personalized medicine. Cybercriminals attack all types of data repositories on a daily basis in an attempt to steal valuable information. Recent estimates indicate that a large company may have to fend off upwards of 1,400 cyber-attacks per week (telecomasia.net, 2013). Medical data breaches and medical identity theft are growing target areas for cyber criminals. Besides mitigating network attacks, a cloud vendor can also help maintain the privacy of a customer's patient data. This can be accomplished through advanced firewall techniques and high-level network encryption. These techniques help ensure that the data on their servers is only viewed by those with the proper clearance. However, these need to be implemented along with proper access control protocols. Appropriate authentication and authorization of users must occur. Authentication verifies that the user is who they say they are; and authorization verifies that the user has the correct permissions for the tasks they wish to complete.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

When personalized medicine clients are using cloud computing, they are likely to be storing sensitive patient information, even though this data would have been stripped of any personal identifiers. The social and legal ramifications of unauthorized disclosure of healthcare data outsourced to the cloud is of concern to both individuals and governments (Wang, Chen, & Zhang, 2015). A recent failed implementation of a cloud solution for a hospital cited data security risks as one of the primary concerns of healthcare providers (Sultan, 2014). The vendor must ensure that the data is neither lost nor stolen. Both private and public cloud solutions are vulnerable to cyber security breaches. A recent breach into personal iCloud accounts highlights the multiple vulnerable points that can exist with web-based cloud computing (Crosman, 2014). Information security and privacy is paramount for success of cloud computing for personalized medicine.

5.3. Contractual Obligations

Thirdly, contractual obligations for services can be complex, difficult and even misleading. On an individual level, problems seem to arise frequently between customers and Internet service providers, mobile phone companies, and cable companies. Grievances usually are made because customers aren't receiving services promised or are being charged more than anticipated. These same problems, and more, can arise with cloud computing. A cloud-computing client will likely have different levels of user privileges and needs, all of which need to be properly handled by the cloud vendor in the personalized medicine agreement.

For the personalized medicine cloud computing client, clear delineation of the relationships among all parties must occur. This includes clarification on items such as costs, duration of services, and types of services provided. Companies need to maintain comprehensive records pertaining to cloud contracts as well. Recently the Department of Defense (DOD) was unable to successfully complete a cost benefit analysis of their cloud solutions, because there was insufficient documentation about cloud contracts (Fabey, 2016). Given the many options for cloud computing including IaaS, PaaS, SaaS and others, clarity of contractual obligations cannot be over-emphasized.

5.4. Regulatory Compliance

The fourth risk to using cloud computing for personalized medicine is government regulation. In the United States, the main federal acts relevant to cloud computing and personalized medicine are HIPAA (1996), GINA (2008), and HITECH (2009). The Health Information Portability and Accountability Act (HIPAA) of 1996 governs the privacy and security of healthcare data. HIPAA mainly addresses how confidential information is protected by healthcare providers, and how that information is shared. Even with a 2013 Omnibus Final Rule update to HIPAA, it still does not directly address cloud storage of patient health information (PHI) (Chang 2013-2014). However, the term "business associates" is used in the omnibus rule which can be interpreted to include cloud vendors (Shwayri, 2014).

Further, neither does HIPAA specifically deal with the nuances of data breaches related to genetic material. The fear is that genetic discrimination, the mal-use of an individual's genetic material in an illegal or unethical manner, can occur and cause harm to individuals. To address this concern, the Genetic Information Non-discrimination Act (GINA), which prohibits insurance companies and employers from discriminating against individuals based on their genetic profile, was passed into law in 2008. The passage of GINA was aimed at reducing patient fears about genetic testing. However, there are limitations

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in its scope since it does not cover long term care, life insurance, disability insurance; and members of the military and persons under Federal Employees Health Benefits Plans or Indian Health Services (Lea, 2009).

The Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 was signed into law to promote the use of health information technology through incentive programs and strengthen civil and criminal penalties as defined by HIPAA (hhs.gov, n. d.). Healthcare professionals and institutions receive incentive payments from Medicare and Medicaid if they are compliant with the two phases of the HITECH Act: 1. meaningful use of electronic health records (EHRs) by 2015 and 2. provide end-users access to a patient portal (Carlson, 2013). Stage three of the HITECH Act is due to begin in 2017 and focuses on improved outcomes (Hitech Answers, 2016). Health status measures, healthcare population patterns and provider collaborations are paramount to phase three of the HITECH Act for improved outcomes (2010). Herein lies an opportunity for personalized medicine with cloud computing by integrating data and using technology to improve healthcare outcomes. However, timely compliance with regulatory requirements present both financial and managerial hurdles for some healthcare providers and other related entities.

5.5. Portability

A fifth risk for personalized medicine clients is the portability of services and data between cloud providers (LeManna, 2012). Portability issues of using cloud computing for personalized medicine can be even more severe if vendors are using specialized medical applications that would restrict clients from moving to other vendors where those software products are not available. Portability refers to both the movement of services from one cloud to another, as well as the movement of data from desktops to a cloud environment (Chang et al., 2013).

When a client adopts a vendor's software or a vendor develops software to meet a client's needs, the software systems will likely be proprietary. Over time the personalized medicine client(s) may decide to transfer data and or services to other cloud providers. Cloud portability is relevant regardless of the cloud services provided i.e. PaaS, IaaS, or SaaS (Chang et al., 2013). Difficulty may arise during this process because of the incompatibility between systems. Additionally, a client may have difficulty integrating the cloud vendor's software into their existing medical software suite. The term interoperability defines the ability of one system to communicate/integrate seamlessly with another system. Interoperability issues should be minimized if changes in cloud services occur.

6. OPPORTUNITIES AND RECOMMENDATIONS

The first research question addressed in this paper focused on the challenges associated with using cloud computing in the area of personalized medicine. This has been extensively examined in the preceding sections of the paper. In this section the second question, how cloud computing can be of value to personalized medicine is addressed in more specific detail. In spite of the many challenges that exist, technology poses many opportunities for the field of healthcare (Chow-White et al., 2015). Personalized medicine is one node under the vast universe of the topic of healthcare.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Our recommendation is that based on the challenges discussed in the preceding section, the client must evaluate many different factors to find the appropriate cloud computing solution based on their medical goals. These two solutions will address many of the security and privacy challenges that dominate the cloud in healthcare conversation. The Translational Genomics Research Institute (TGEN), a non-profit consortium, uses a private cloud supported by Dell to process human genomic data (Borzykowski, 2015).

According to Wang and colleagues (W. Wang et al., 2015) high dimensional medical data can include over 100 variables with data points such as blood pressure, heart rate and others. This type of profile is consistent with the high dimensional data that is part of the personalized medicine profile. Wang et al. (2015) further addressed security concerns of healthcare data on the cloud by using a hybrid cloud with a privacy-preserving protocol that moves sensitive data to the private cloud and less sensitive data to the public cloud.

Companies are looking at creative cloud solutions for personalized medicine initiatives. Through computational mining of multiple large data sets, it is possible to re-identify individuals that were previously part of anonymized genome data (Chow-White et al., 2015). One option is to create a virtual private cloud hosted on a public domain (Sultan, 2014). Ultimately the final decisions are based on evaluating the five main variables: service quality; security and privacy; contractual obligations; government regulations; and portability.

Public sector participation through National Institutes of Health (NIH) and the Food and Drug Administration (FDA) has made funding available to pursue some personalized medicine projects (Hamburg & Collins, 2010). Many other collaborative opportunities in personalized medicine exists both through public sector and private entities. Growth in the area of personalized medicine can also lead to more drug clinical trials based on the genetic understanding of disease (Hamburg & Collins, 2010). Ultimately, using the cloud for personalized medicine is an ongoing journey with many opportunities for growth.

7. CONCLUSION

The fields of medicine and ICTs are not strangers to each other. In fact, there is a healthy symbiotic relationship between these two giants as they equally strive to improve the quality of human life. The area of cloud computing for personalized medicine lends itself as stellar example for the intersection of medicine and technology. Personalized medicine is an important area in the field of medicine, and there are many opportunities to innovate through technology. Cloud computing presents many powerful advantages to bring disparate entities together to advance the field of personalized medicine. In spite of the apparent risks associated with both cloud computing and personalized medicine, the value and future potential of this marriage is too significant to ignore. In this paper we bring this issue to the foreground and discuss both the value and challenges of using cloud computing for personalized medicine.

REFERENCES

Ardagna, D., Casale, G., Ciavotta, M., Pérez, J. F., & Wang, W. (2014). Quality-of-service in Cloud Computing: Modeling Techniques and their Applications. *Journal of Internet Services and Applications*, *5*(11), 2–17.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Arlotto, P. N. (2010). 7 Strategies for Improving HITECH Readiness. *Healthcare Financial Management, 64*(11), 90-96.

Badger, L., Grance, T., Patt-Corner, R., & Voas, J. (2012). *Special Publication 800-146 Cloud Computing Synopsis and Recommendations: National Institute of Standards and Technology*. NIST.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., ... Garraway, L. A. (2012). The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, *483*(7391), 603–607. doi:10.1038/nature11003 PMID:22460905

Bendor-Samuel, P. (2014). Personalized Medicine Rests On Cloud At H.D. Smith. Retrieved from http://www.informationweek.com/strategic-cio/digital-business/personalized-medicine-rests-on-cloud-at-hd-smith/a/d-id/1315726

Bolouri, H. (2010). Introduction. In *Personalized Genomics and Personalized Medicine* (pp. 1–18). Imperial College Press.

Borzykowski, B. (2015). Decoding Your 'Instruction Manual': Inside The Amazing Advances In Personalized Medicine. *Forbes.com*. Retrieved from http://www.forbes.com/sites/dell/2015/07/28/decoding-your-instruction-manual-inside-the-amazing-advances-in-personalized-medicine/

Carlson, J. (2013). Healthcare goes Hitech. *Baylor Business Review*, *32*(1), 28–31.

CDC. (2013). Deaths and Mortality. Retrieved from http://www.cdc.gov/nchs/fastats/deaths.htm

Center for Biomedical Informatics (CBMI) at Harvard Medical School. (2014). Laboratory for Personalized Medicine. Retrieved from http://lpm.hms.harvard.edu/

Chang, V., Walters, R. J., & Wills, G. (2013). The development that leads to the Cloud Computing Business Framework. *International Journal of Information Management*, *33*(3), 524–538. doi:10.1016/j.ijinfomgt.2013.01.005

Chang (2013-2014). The Dark Cloud of Convenience: How the new HIPAA Omnibus Rules Fail to Protext Electronic Personal Health Information. *Loyola of Los Angeles Entertainment Law Review, 34*(2), 119-154.

Chow-White, P. A., MacAulay, M., Charters, A., & Chow, P. (2015). From the bench to the bedside in the big data age: Ethics and practices of consent and privacy for clinical genomics and personalized medicine. *Ethics and Information Technology*, *17*(3), 189–200. doi:10.100710676-015-9373-x

Crosman, P. (2014). Cloud Security: Trouble in the Clouds. *Bank Technology News, 27*(3).

Deakin, D. (2011). Personalized Medicine - How will the end of one-size-fits-all medicine effect the clinical trial industry?

Fabey, M. (2016). IG Faults Pentagon Cloud Computing Service Contracts. *Aerospace Daily & Defense Report*, *255*(3), 5.

Forrest, C. (2016). Amazon Unveils New Data Center Hardware to Get You Securely into the Cloud. *TechRepublic*. Retrieved from http://www.techrepublic.com/article/amazon-unveils-new-data-center-hardware-to-get-you-securely-into-the-cloud/

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gray, B. (1989). *Collaborating: finding common ground for multiparty problems*.

Hamburg, M. A., & Collins, F. S. (2010). The Path to Personalized Medicine. *The New England Journal of Medicine*, *363*(4), 301–304. doi:10.1056/NEJMp1006304 PMID:20551152

hhs.gov. (n. d.). HITECH Act Enforcement Interim Final Rule. Retrieved from http://www.hhs.gov/hipaa/for-professionals/special-topics/HITECH-act-enforcement-interim-final-rule/index.html

Hitech Answers. (2016). The Stages of Meaningful Use. Retrieved from http://www.hitechanswers.net/ehr-adoption-2/meaningful-use/

Horowitz, B. T. (2014). Elderly Patients At Risk From Drug Side Effects. *Interaction*, *14*(May). Retrieved from http://www.cruxialcio.com/researchers-enlist-ibm-cloud-quest-better-safer-elder-care-7476

Jayanthi, A. (2014). 10 Biggest Technological Advancements for Healthcare in the Last Decade. *Becker's Health IT & CIO Review*. Retrieved from http://www.beckershospitalreview.com/healthcare-information-technology/10-biggest-technological-advancements-for-healthcare-in-the-last-decade.html

Johnson, A. (2016). An Impediment to Personalized Medicine Commercialization is the Lack of Understanding of the Value of the Testing. *Journal of Commercial Biotechnology*, *22*(1), 3–5. doi:10.5912/jcb732

Lea, D. H. (2009). The Genetic Information Nondiscrimination Act (GINA): What it Means for Your Patients and Families. *Journal*, *14*(2). Retrieved from http://www.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAPeriodicals/OJIN/TableofContents/Vol142009/No2May09/Articles-Previous-Topics/The-Genetic-Information-Nondiscrimination-Act-GINA.html

LeManna, L. (2012). *Top 9 Challenges in Cloud Computing* (Vol. 2013). SAP Business Innovation.

Martin-Flatin, J. P. (2014). Challenges in Cloud Management. *IEEE Cloud Computing*, *1*(1), 66–70. doi:10.1109/MCC.2014.4

Neu, S. C., Crawford, K. L., & Toga, A. W. (2016). Sharing data in the global Alzheimer's association interactive network. *NeuroImage*, *124*, 1168–1174. doi:10.1016/j.neuroimage.2015.05.082 PMID:26049147

Podshivalov, L., Fischer, A., & Bar-Yoseph, P. Z. (2014). On the Road to Personalized Medicine: Multiscale Computational Modeling of Bone Tissue. *Archives of Computational Methods in Engineering*, *21*(4), 399–479. doi:10.100711831-014-9120-1

Puustjärvi, J., & Puustjärvi, L. (2013). Practising Cloud-based Telemedicine in Developing Countries. *International Journal of Electronic Healthcare*, *7*(3), 181–204. doi:10.1504/IJEH.2013.057407 PMID:24191340

Rehm, H. L., Berg, J. S., Brooks, L. D., Bustamante, C. D., Evans, J. P., Landrum, M. J., ... Watson, M. S. (2015). ClinGen — The Clinical Genome Resource. *The New England Journal of Medicine*, *372*(23), 2235–2242. doi:10.1056/NEJMsr1406261 PMID:26014595

Samba, A. (2012). Logical Data Models for Cloud Computing Architectures. *IT Professional*, *14*(1), 19–26. doi:10.1109/MITP.2011.113

Sheng, J., Li, F., & Wong, S. T. C. (2015). Optimal Drug Prediction From Personal Genomics Profiles. *IEEE Journal of Biomedical and Health Informatics, 19*(4).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shwayri, R. N. (2014). Balancing the Risks and Rewards of Cloud-Based Healthcare Information. *Information & Management*, *48*(3), 42–45, 47.

Sultan, N. (2014). Making use of cloud computing for healthcare provision: Opportunities and challenges. *International Journal of Information Management*, *34*(2), 177–184. doi:10.1016/j.ijinfomgt.2013.12.011

telecomasia.net. (2013). Akamai, IBM team for cloud security. Retrieved from http://www.telecomasia.net/content/akamai-ibm-team-cloud-security

U.S. Food and Drug Administration FDA (2013). Paving the Way for Personalized Medicine. Retrieved from http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/PersonalizedMedicine/UCM372421.pdf

U.S. National Library of Medicine. (2015). NIH Data Sharing Repositories. Retrieved from https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html

Viceconti, M., Hunter, P., & Hose, R. (2015). Big Data, Big Knowledge: Big Data for Personalized Healthcare. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1209–1215. doi:10.1109/JBHI.2015.2406883 PMID:26218867

Wang, D. (2015). What will cloud computing bring to healthcare? – Overview of cloud applications in healthcare. *International Journal of Healthcare Technology and Management*, *15*(1), 49–57. doi:10.1504/IJHTM.2015.070540

Wang, W., Chen, L., & Zhang, Q. (2015). Outsourcing high-dimensional healthcare data to cloud with personalized privacy preservation. *Computer Networks*, *88*, 136–148. doi:10.1016/j.comnet.2015.06.014

Wood, L. (2015, June 9). Cloud market in healthcare technologies. *Business and Markets*.

Yang, H., & Tate, M. (2012). A Descriptive Literature Review and Classification of Cloud Computing Research. *Communications of the Association for Information Systems*, *31*, 35–60.

Yang, W., Soares, J., Greninger, P., Edelman, E., Lightfoot, H., Forbes, S., ... Garnett, M. J. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Research*, *14*(D1), D955–D961. doi:10.1093/nar/gks1111 PMID:23180760

This research was previously published in the International Journal of E-Health and Medical Communications (IJEHMC), 8(2); edited by Joel J.P.C. Rodrigues; pages 1-17; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 69
Analysis of Healthcare Workflows in Accordance with Access Control Policies

Sandeep Lakaraju
Dakota State University, USA

Dianxiang Xu
Boise State University, USA

Yong Wang
Dakota State University, USA

ABSTRACT

Healthcare information systems deal with sensitive data across complex workflows. They often allow various stakeholders from different environments to access data across organizational boundaries. This elevates the risk of exposing sensitive healthcare information to unauthorized personnel, leading 'controlling access to resources' a major concern. To prevent unwanted access to sensitive information, healthcare organizations need to adopt effective workflows and access control mechanisms. Many healthcare organizations are not yet considering or do not know how to accommodate the 'context' element as a crucial element in their workflows and access control policies. The authors envision the future of healthcare where 'context' will be considered as a crucial element. They can accommodate context through a new element 'environment' in workflows, and can accommodate context in policies through well-known attribute based access control mechanism (ABAC). This research mainly addresses these problems by proposing a model to integrate workflows and access control policies and thereby identifying workflow activities that are not being protected by access control policies and improving the workflow activities and/or existing access control policies using SARE (Subject, Action, Resource, and environment) elements.

DOI: 10.4018/978-1-7998-1204-3.ch069

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The wide adoption of information technology across healthcare organizations has increased the need for developing efficient workflows. Literature has shown that healthcare organizations with efficient workflows have more success rate in adopting healthcare information technology (HIT) into their organizations (Sittig, Krall, Kaalaas-Sittig, & Ash, 2005). Effective care delivery is possible through workflows with competent activities (Campbell, Sittig, Ash, Guappone, & Dykstra, 2006). Nevertheless, security and privacy of healthcare information have become major challenges for healthcare organizations (Akinyele et al., 2011; Alhaqbani & Fidge, 2007). In particular, effective access control mechanisms are needed to protect sensitive information from being exposed to unauthorized personnel.

Workflows

Workflows are used to describe the pattern of tasks to be executed by users to achieve business objectives. According to Welch (2014), "Efficient clinical workflow saves time, saves money, and saves lives. And in today's industry, workflow can have a significant effect on reimbursement". Workflows can be very complicated, especially in a complex environment like healthcare which may involve various subjects trying to perform actions on certain resources in multiple environments, thus requires controlling the access of resources by subjects. For a subject to perform an action on a resource in an environment, that subject should be authorized to perform the intended action. In this research, Business Process Modeling Notation (BPMN) has been used to develop workflows. BPMN (Business Process Modeling Notation) is visualization for business process workflows (Giaglis, 2001). The various elements of BPMN are given in Figure 1. We use four types of elements, defined as follows: Events that can be start or finish, distinguishable by size of circles; Activities (Tasks) have an "id" (A1; A2 etc.) and possibly user icons on the upper right corner of the rectangle, and groups indicate a group of activities; Gateways are diamond shaped, with multiple input channels; Conditional gateways are plain diamonds with only one output channel that is based on the evaluation of a condition; Parallel gateways have diamonds with a '+' sign inside. They have multiple input/output channels indicating that control flows on those channels in parallel; Sequences (flows) link tasks together, and associations are used to associate activities with flow objects; some of the extra events used are also shown in the Figure 1 representing Binding of Duties (BOD) policy and Separation of Duty (SOD) policy, and a release event.

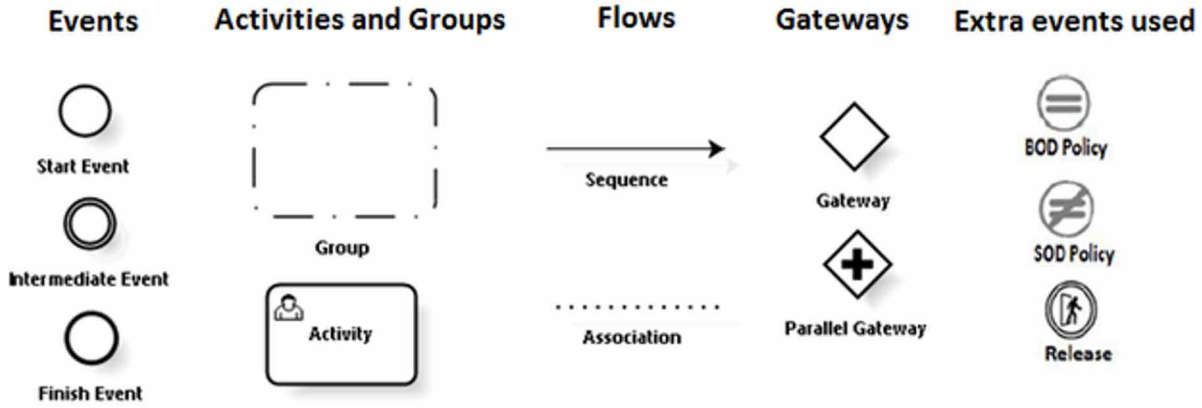
A workflow involves execution of a series of certain activities which help in achieving a goal (Chaari, Biennier, Amar, & Favrel, 2004). These workflows have to be executed in a secure way, which is made possible through access control.

Access Control

Inter-organizational systems allow users to access and share data beyond organizational boundaries and therefore needs proper authorization mechanisms to protect sensitive information from being exposed to unauthorized personnel. An access control policy defines the conditions to which access to resources can be granted and to whom (Ferreira, Cruz-Correia, Antunes, & Chadwick, 2007). With the increasing complexity of information systems, access control methods have evolved from Mandatory Access

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Workflow elements



Control (MAC), Discretionary Access Control (DAC), Task Based Access Control (TBAC), Context Based Access Control (CBAC), Role-Based Access Control (RBAC), to Attribute-Based Access Control (ABAC) (Xu & Zhang, 2014).

Attribute Based Access Control (ABAC)

ABAC provides dynamic access control capability. ABAC does not depend on subject-resource relation, but on various attributes (of subject, resource, environment, etc.). National Institute of Standards and Technology (NIST) indicates ABAC as a recommended access control language which can improve 'information sharing within organizations and between diverse and disparate organizations while maintaining control of that information' (Hu, Scarfone, & Kuhn, 2013).

Coming to attribute base access control (ABAC), it is "an access control method where subject requests to perform operations on objects are granted or denied based on assigned attributes of the subject, assigned attributes of the object, environmental conditions, and a set of policies that are specified in terms of those attributes and conditions" (Hu et al., 2013). The preferability of ABAC over other access control mechanisms is demonstrated in detail in (Yuan & Tong, 2005). In ABAC, the authorization elements are defined in terms of attributes, rather than identities, of subjects, actions, resources, and environments (SARE). Attributes are characteristics of these entities that are predefined and pre-assigned by an authority. Subject attributes help in identifying a user (or an operation or a process) with its characteristics rather than its role; for example: job title, user name, age, etc. A resource attribute is a characteristic of a resource (record, data, etc.); for example: medical elements of patient record, non-medical elements of patient record, etc. Action attributes help in identifying the type of action the subject will perform on a resource; for example: read, write, etc. Environment attributes are operational or situational characteristics, such as current time and IP address. Consider the following access control rule: Any doctor can read and write into patient health record (PHR) only for his/her designated patients from internal network. The subject attributes include 'doctor', the action attribute is 'read' and 'write', and the resource attribute is 'patient health record', and the environment variable is 'internal network' and the condition is that the doctor can only access the health records of the patients designated to him/her.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

HIPAA Compliance of Workflows and Access Control

Implementing HIPAA compliance involves the steps below: (i) establishing a process to develop/improve policies (ii) administering security functions (iii) actively tracking personnel changes, and (iv) examining and recording the policy application (K. Cole, 2002).

The privacy rule involves: (i) Minimum necessary use: where an organization should delineate people who may request to have their hands on PHI. Organizations should also delineate the health information into various categories and should define the level of access an employee should have to access that particular information, and any other specific regulations that may apply; (ii) notice of privacy practice: Organizations should make understand the employees and any other stakeholders about their privacy practices, (iii) it also provides the patients with advanced rights of access about their health information. Upon patient's request, a healthcare entity should provide information sharing policies, and patients can request with whom their information may be shared or not shared.

The security rule involves confidentiality, integrity, and availability of information. Confidentiality involves restriction of access to information. Integrity imposes strict restrictions on type of modifications that can be performed on the protected health information. Availability ensures that recognized users have access to PHI whenever they require. This can be implemented in three ways as shown in this research: developing policies in plain English, implementing through computer models (such as workflows and ABAC), and using technical mechanisms. In this research, the access control policies and workflow activities are HIPAA compliant because:

1. This methodology requires to develop and use uniform SARE standards
2. Workflow activities and access control policies use same SARE attributes
3. Access control policies consider environment variables to allow or deny access
4. Policies can be modified based on patient's interests
5. The privacy will be enhanced by considering the environment attributes, which would further allow the administrator to know what information has been accessed from which location and by whom.

Integrating Access Control and Workflow

Many organizations treat access control and workflow as two separate issues. As both workflow specifications and access control polices can be complex, it is likely that there exists discrepancy between workflow activities and access control policies. For example, some workflow activities may not be adequately regulated by the access control policies. To address this issue, we propose a methodology for integrated modeling and analysis of workflows and ABAC policies.

The proposed methodology views workflow activities in terms of subject, resource, action, and environment. Specifying environment attributes in a workflow will provide more contextual information for analysis. It is also consistent with the ABAC concepts. Through cross-examination of inter-organizational workflow activities and access control policies, the proposed methodology can identify: (a) Whether there are workflow activities that are not covered by the access control policies, (b) whether there are unused access control policies, (c) how many policies are covering an activity, and (d) whether the workflow activities and the access control policies need to be modified or refined.

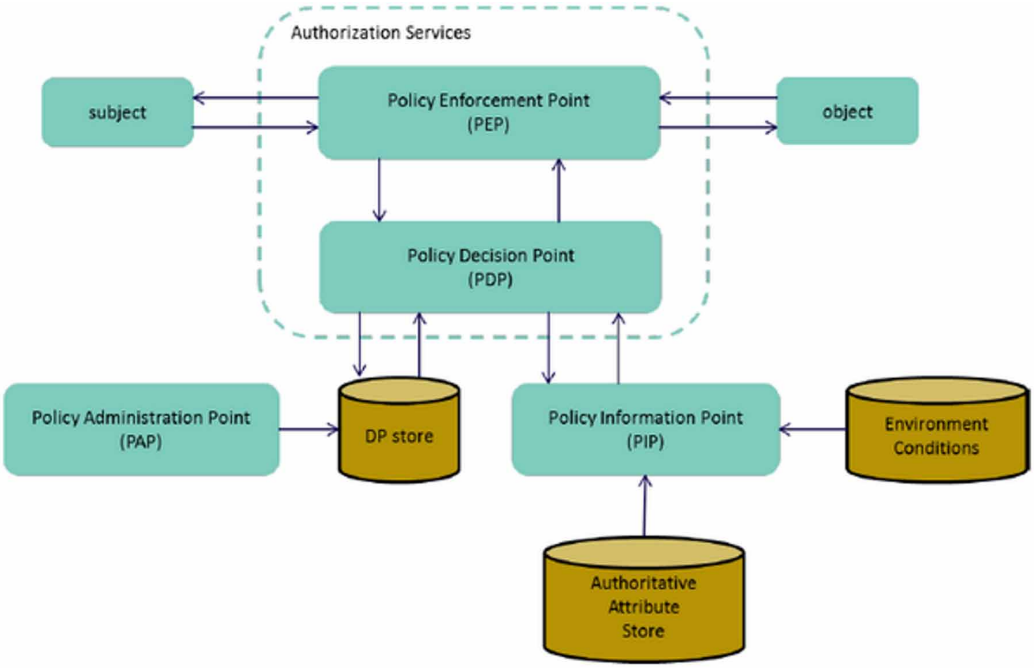
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Organizational Implementation of ABAC

An access control policy defines the conditions to which access to resources can be granted and to whom (Ferreira et al., 2007). With the increasing complexity of information systems, access control methods have evolved from Mandatory Access Control (MAC), Discretionary Access Control MAC (DAC), Access Control List (ACL), Role-Based Access Control (RBAC), to Attribute-Based Access Control (ABAC) (Xu & Zhang, 2014). The preferability of ABAC over other access control mechanisms is demonstrated in detail in (Yuan & Tong, 2005). In ABAC, the authorization elements are defined in terms of attributes, rather than identities, of subjects, actions, resources, and environments (SARE). Attributes are characteristics of these entities that are predefined and pre-assigned by an authority. A sample organizational ABAC implementation has been shown in Figure 2. It mainly involves certain functional 'points': Policy Enforcement Point (PEP), Policy Decision Point (PDP), Policy Administration Point (PDP), and Policy Information Point (PIP). Considering a particular scenario where a subject requests access to a resource under certain environment, this is sent as a request into a policy enforcement point (PEP) where a decision is taken whether to 'allow' that access or 'deny' the access based on certain attributes (user, action, resource, and environment). PEP sends that request to PDP where the request is broken into subject attributes, resource attributes, action attributes and environmental attributes.

Initially, each of those attributes is stored in their respective attribute stores managed at the policy information point (PIP). And the Policy Administration Point (PAP) is the point where policies are managed by the administrators. PDP will validate the request based on the available policies from PAP and the attributes available from PIP. Once the PDP is able to validate the request, it will send a notification

Figure 2. Access control mechanism distribution in a sample organizational ABAC implementation (Hu et al., 2013)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to PEP suggesting to 'allow' the access, otherwise 'denies' the request. To support ABAC, OASIS has developed the eXtensible Access Control Markup Language (XACML) standard (Identifier, 2005). Consider the following access control rule: *Any doctor can read and write into patient health record (PHR) only for his/her designated patients from internal network*. The subject attributes include 'doctor', the action attribute is 'read' and 'write', and the resource attribute is 'patient health record', and the environment variable is 'internal network' and the condition is that the doctor can only access the health records of the patients designated to him/her.

Literature Review

A workflow involves execution of a series of certain activities which help in achieving a goal (Chaari et al., 2004). These workflows must be executed in a secure way, which is possible through access control. To achieve these secured workflows, various access control techniques have been studied in accordance with the workflows, but none of the works have shown an integrated analysis and have only dealt with workflows and access controls separately. In (Chaari et al., 2004), authors have extended role based access control and proposed an authorization and access control model for workflows. Research by Chaari et al. (2004) mainly focuses on the user's role and the task the user has to execute; but it does not consider any other contextual information or user attributes except the role. Russello et al. (2008) have proposed a workflow based access control framework for e-health applications, where permission is granted to a user based on the necessity of the task that the users have to execute. Russello et al. (2008)'s research mainly focuses on the 'need' of the user to execute the task, but does not consider any of the environment or resource or subject attributes except the role. Russello et al. (2008) indicates that 'access rights must be granted only to entities that require access to a given resource and just for the amount of time that access is necessary', but healthcare is rather dynamic environment where access is required in varying contexts which requires precise contextual information to allow a user to access certain resource. This research will address this issue by focusing on importance of 'environment' attributes and our methodology will demonstrate the integrated analysis of workflow and access control which can grant access rights only to the entities that require access to a given resource under certain context where access is necessary. Le, Doll, Barbosu, Luque, & Wang (2012) have improved the traditional role based access control model to enhance the information access management in the environment of workflow and team collaboration in healthcare programs. Le et al. (2012)'s research has mainly concentrated on traditional RBAC and is an extension of RBAC. This research deals with ABAC, which not only deals with user attributes such as 'role' but also with various resource, environment and action attributes. Lu, Zhang, & Sun (2009) have developed a framework to specify security constraints in workflow by improving business process management notation model. Authors have improved the workflow model to aid role based access, separation of duty, history based task assignment, etc. Compared to this work, our model deals with the attributes of the subjects, action, resource, and environments rather than role based or SOD or history based or task based. Zhang & Liu (2011) have developed a workflow oriented attribute based access control model that provides 'Service-Oriented Computing', where any device's functionality can be offered as a standard service. Although Zhang & Liu (2011) talk about attribute based access control, they have not considered the need to identify unused policies or the need to identify unprotected workflow activities. XiangPeng, Cerone, & Krishnan (2006) have developed a framework integrating workflows (BPEL) and access control (RBAC) to execute security constraints through 'temporal logic'. Though the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

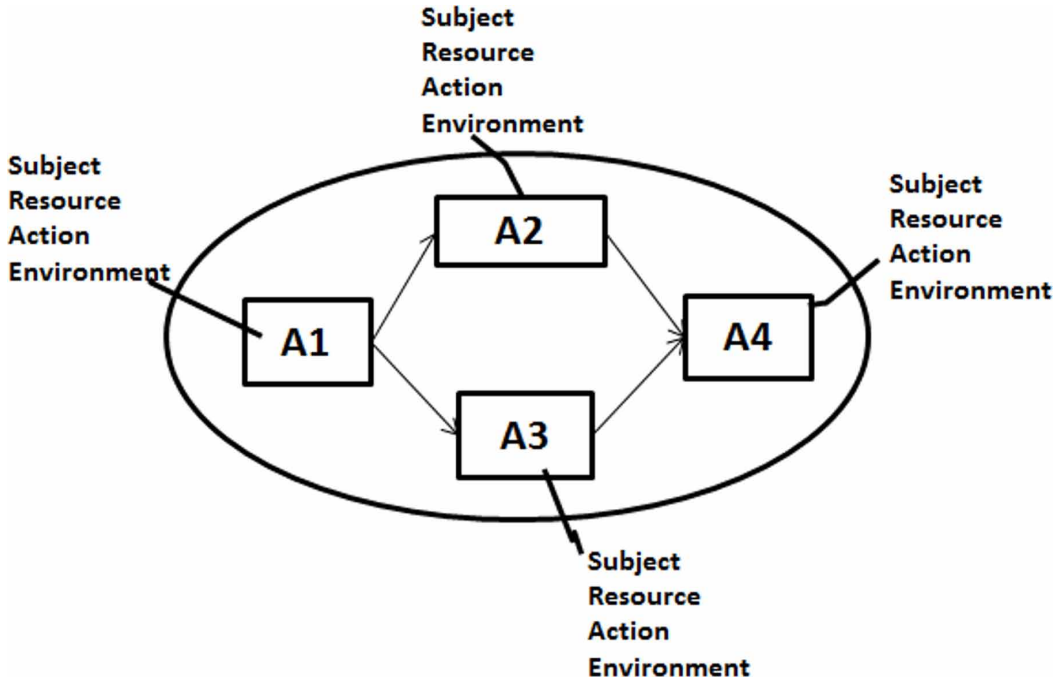
framework deals with integrating workflows and RBAC through temporal logic; regular people may not exercise the logic to execute security restraints. Compared to their work, our model deals with enforcing ABAC policies on the workflow activities which is not covered in (XiangPeng et al., 2006).

None of the above mentioned literature talks about integrated analysis of workflow and ABAC policies. One of the main takeaways from our research is that it is possible to identify any unused policies and also it is possible to identify any workflow activities that are not being protected by access control policies.

Assumptions

Some of the assumptions in developing the current model are discussed below. We are not developing a new modeling language. We are adding a new element 'environment' to the activity of the workflow. We are suggesting breaking down an activity in the workflow in terms of 'Subject, Action, Resource, and Environment' (SARE) elements as shown in Figure 3. Not every activity may have all of the four (SARE) elements, i.e. each activity may have any of the SARE elements. We identify common themes between the workflow and ABAC policies in terms of SARE elements. In any given situation to integrate a workflow and policies, we assume that workflow and policies will have any or all of the SARE elements. We assume workflow is a valid one without any deadlocks. Organization will define the SARE elements to be used commonly by both workflow and the ABAC policies.

Figure 3. Workflow activity breakdown



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Methodology

The general idea of integrated modeling and analysis is as follows: (a) specify workflow activities in terms of SARE elements, (b) specify ABAC policies, (c) cross-examine the workflow activities and ABAC policies to identify discrepancy, (d) if there is no discrepancy, then terminate, otherwise go back to (a) if the discrepancy is about the workflow activities, or go back to (b) if the discrepancy is about the ABAC policies. In the following, we focus on the cross-examination of workflow activities and ABAC policies, as illustrated in Figure 4.

Step 1: The given workflow may include many activities (A1, A2, A3…). Each activity has to be individually identified from the given workflow as (A1), (A2), and (A3) ….

Step 2: From the individual activity that has been identified from the workflow (say A1), subject resource action and environment (SARE) attributes have to be identified, i.e. each activity (A1) has to be broken down into SARE (subject action resource and environment) elements (Figure 4's break the activity in terms of SARE).

Step 3: ABAC policies will be written targeting the Subject and/or Resource elements of the SARE attributes. An ABAC policy is represented by the target which the policy is intended to cover. It is seen in Figure 4 where P1 is represented by R1, where R1 is the resource1 targeting any workflow activity which involves the resource1. All the workflow activities that involves resource1 (R1) will be covered by this policy. Therefore, in this step, the target element from the policy is extracted and used to represent the policy to match with the respective attribute in the workflow activity.

Step 4: Once the workflow activities are broken into SARE elements, and the targets are extracted from the Policies, each of the 'subject and resource' attributes of the workflow activity are cross examined across the available ABAC policies that are targeting the respective subject and/or resource attributes.

Step 5: It is important to identify all the policies that can be applied on an activity. Once the policies are applied on the activity of the workflow, record each of those policies per workflow activity into a table (Table 1). Identify all of the policies that are being applied on each of the activity of the workflow and develop a table (Policies applicable on each activity of the workflow: Table 1).

Step 6: Once the policies are applied on the activity of the workflow, identify all of the policies that are not being used on the workflow (Policies that are not being used on the workflow: Table 2).

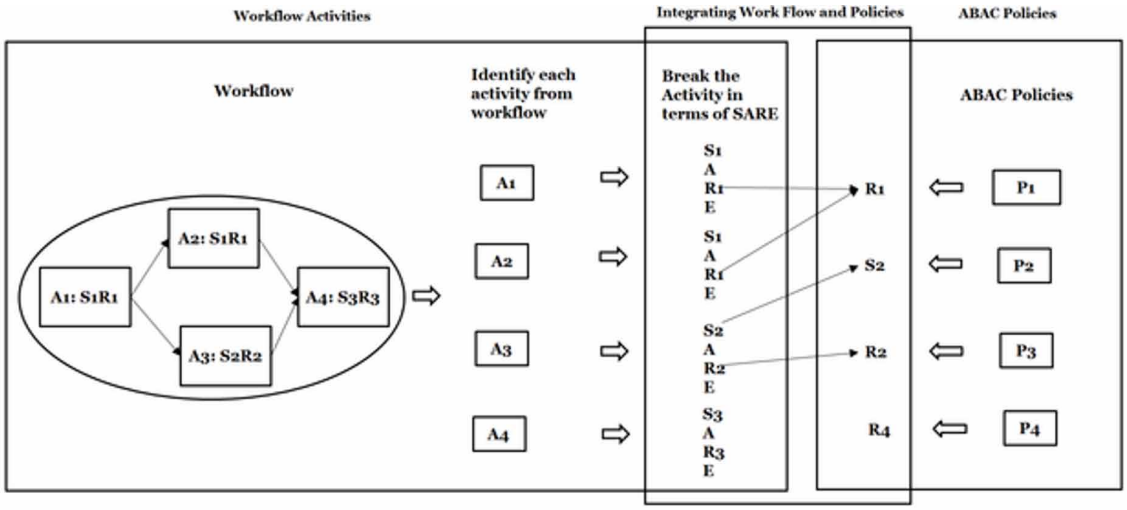
Implementing the Approach

Figure 4 represents the integrated model of the workflow and ABAC policies. It is evident from the Figure 4 that the 'workflow' denotes a workflow with four activities: A1, A2, A3, and A4. Each of these activities has 'subject, resource, action, and environment attributes'. Once there is the workflow, each of the activities of the workflow is extracted from the workflow as shown under the 'Identify each activity from workflow' task in Figure 4 ({A1}, {A2}, {A3}, and {A4}). Following this, break each of these activities into the subject, action, resource, and environment elements (SARE).

{A1} => {S1, A, R1, E}; {A2} => {S1, A, R1, E}; {A3} => {S2, A, R2, E}; and {A4} => {S3, A, R3, E}

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Model to integrate workflow and ABAC policies



Similarly, from the given ABAC policies {P1, P2, P3, and P4}, identify individual policies. Each of these policies will further include multiple rules which are written targeting either a subject or a Resource. Based on the target, each of these policies may be targeted to either subject or resource. Therefore, each of these policies can be shown as under 'ABAC Policies' in Figure 4, P1=>R1, P2=>S2, P3=>R2, P4=>R4. Once the subject and resource elements are found from the workflow activities, the policies with respective subject or resource targets must be imposed on that particular activity as shown below.

As {A1} has subject S1 and Resource R1, any of the policies that target Subject S1 or Resource R1 can be imposed. From the given policies, only P1 is applicable. This case is a special example where there is no policy specified on subject S1, and therefore this activity of the workflow is **incompletely covered (partial)** by the access control policy. It is the same case with A2.

Whereas {A3} has a subject S2 and Resource R2, and therefore, policies P2 and P3 can be applied on the activity A2 showing that A2 is completely covered.

Coming to the activity {A4}, it has subject S3 and Resource R3, but there is no policy targeting either subject S3 or Resource R3. Therefore, this activity is not at all covered. Similarly, Policy P4 is targeting resource R4, which has no occurrence in the workflow, resulting in no use of this policy. The outputs from the current model are shown in the below Tables 1 and 2. Based on these outputs, the workflow developers and access control policy administrators can make the necessary changes to the respective workflows and policies.

Expected Results of the Integration Model

Based on these outputs from Tables 1 and 2, it is evident that (i) Activities A1 and A2 are each covered by one Policy P1, whereas Activity A3 is covered by two policies P2 and P3, (ii) The workflow administrators that the 'IOS workflow activity A4 is not being covered by any ABAC policy and therefore it is recommended to re-evaluate the activity A4, (iii) the Policy P4 is not covering any workflow activity and therefore it needs re-evaluation, (iv) in re-evaluation, policy administrators may re-evaluate the un-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Policies applicable on each activity of the workflow

Activity Name	Policy applicable (Yes/ N0/ Partial)?	If applicable, Policy Name
A1	Partial	P1
A2	Partial	P1
A3	Yes	P2, P3
A4	No	-

Table 2. Policies in use (Y/N)

Policy	Policy in use (Y/N)?
P1	Y
P2	Y
P3	Y
P4	N

covered workflow activity and the policies that are not in use, and based on the need, they may suggest the organizations to add new SARE elements using which the administrators may develop new policies or workflow activities.

Demonstrating the Methodology in a Healthcare Scenario

Manipal hospital is located in Bangalore, India with 54 bed count and can accommodate up to 91 beds when fully equipped. It is located in the center of the city. The hospital has six departments: administration department, inpatient/outpatient department, pharmacy department, laboratory department, emergency department, and surgical ward. The hospital accommodates both inpatients and outpatients. The pharmacy department provides medications for both the inpatients and outpatients. The scientific staff of the hospital include: front desk specialists, doctors, nurse specialists, lab specialists, pharmacists, and surgical specialists. This hospital has recently implemented and mandated to use electronic health records across the hospital. The main parts of the electronic health records may include: administrative record, clinical record, pharmacist record, patient record, and surgical record. Administrative record of the EHR mainly includes patient demographic information (such as age, sex, first and last names, email addresses, physical addresses, etc.), date of admission and discharge, patient registration information, etc. This administrative record is restricted to be accessed only by the front desk personnel, nurse specialists and doctors. The clinical record stores the health related information of the patient. The clinical record may include detailed information of pathology reports, blood reports, ENT reports, doctor notes, nurse notes, prescriptions, etc. The pharmacist record includes prescription ids, patient ids, physician ids, and pharmacist notes. The surgical record of the EHR includes the surgical information, images (like Xrays, etc.), surgery specialist, and patient information. The subject, action, resource, and environment (SARE) attributes are provided by the hospital and are discussed below.

Subject attributes may include: department id (which includes an administration department, pharmacy department, emergency department, etc.); specialty id (which includes doctors, nurses, pharmacists, etc.); user id (which includes patient ids, doctor ids, nurse ids, etc.). Action attributes may include: read, write, update, delete, fill, validate, approve, reject, etc.; resource attributes include: electronic health record, patient health record, pharmacist record, clinical record, administrative record, surgical record, etc.; environment attributes include: ip addresses (92.92.1.1, 93.93.1.1, etc.), inpatient ward, outpatient ward, laboratory, surgery ward, doctor office, etc. Each of the departments of the hospital is assigned with different IP address as mentioned below: for hospital administration department: 92.92.1.1; for inpatient

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and outpatient wards, doctors' offices, and nurse quarters, the IP range is from 91.91.1.1 to 91.91.1.5; the internet protocol address for pharmacy department is 93.93.1.1, and for laboratory department is 94.94.1.1, for emergency department is 95.95.1.1, and for surgical ward is 91.91.1.6.

This case study contains multiple workflows which are developed for various causes. A basic scenario in a hospital starts when a patient walks in for an appointment with the doctor (Figure 5). This workflow mainly involves five workflow activities. Subjects involved in this workflow are patient and front desk person (FDP). FDP uses a system which has been configured for the IP address 92.92.1.1 and the FDP is allowed only to use this system from this IP address. The FDP cannot access any other systems configured with other IP addresses. Also the FDP can only access administrative record but not any other clinical information of the patient's electronic health record. When a patient walks in, the front desk person (FDP) asks the patient for his/her demographic information and tries to retrieve the patient information, if the patient is coming in for the first time, the FDP will request the patient to fill in the registration form and FDP validates the form, and thereby registers the patient into the hospital system followed by scheduling an appointment with a doctor.

This particular scenario can be considered as an example for separation of duties (SOD) and also binding of duties (BOD). There are mainly two roles: patient and FDP. A patient will check for an appointment with the FDP. FDP will check if the patient is valid, and then checks for doctor's schedule, and schedules an appointment. For this example, separation of duties applies between patient and receptionist. Main assumptions for a given scenario of patient's appointment with doctor to be static SOD are: (i) receptionist cannot be the patient; (ii) patient himself/herself cannot act as receptionist. If a receptionist is a patient, the receptionist cannot perform tasks checking the validity of the patient, checking for doctor's schedule and scheduling an appointment. Only another receptionist can perform those tasks.

As shown in the Figure 6, subject attributes to activity allocation can be seen, where multiple users with usernames (such as Tresa, Mary, etc) are allocated to subject attributes (role) and are thereby connected to their respective workflow activities.

As shown in the Figure 6, it is evident that activities 'FDP requests for patient demographic information...' and 'Patient fills the patient registration form...' cannot be completed by the same subject, accommodating separation of duty constraint. Similarly, activity 'FDP validates patient info in the admin record…' and activity 'FDP schedules appointment time with doctor…' should be done by the same subject, accommodating binding of duties constraint.

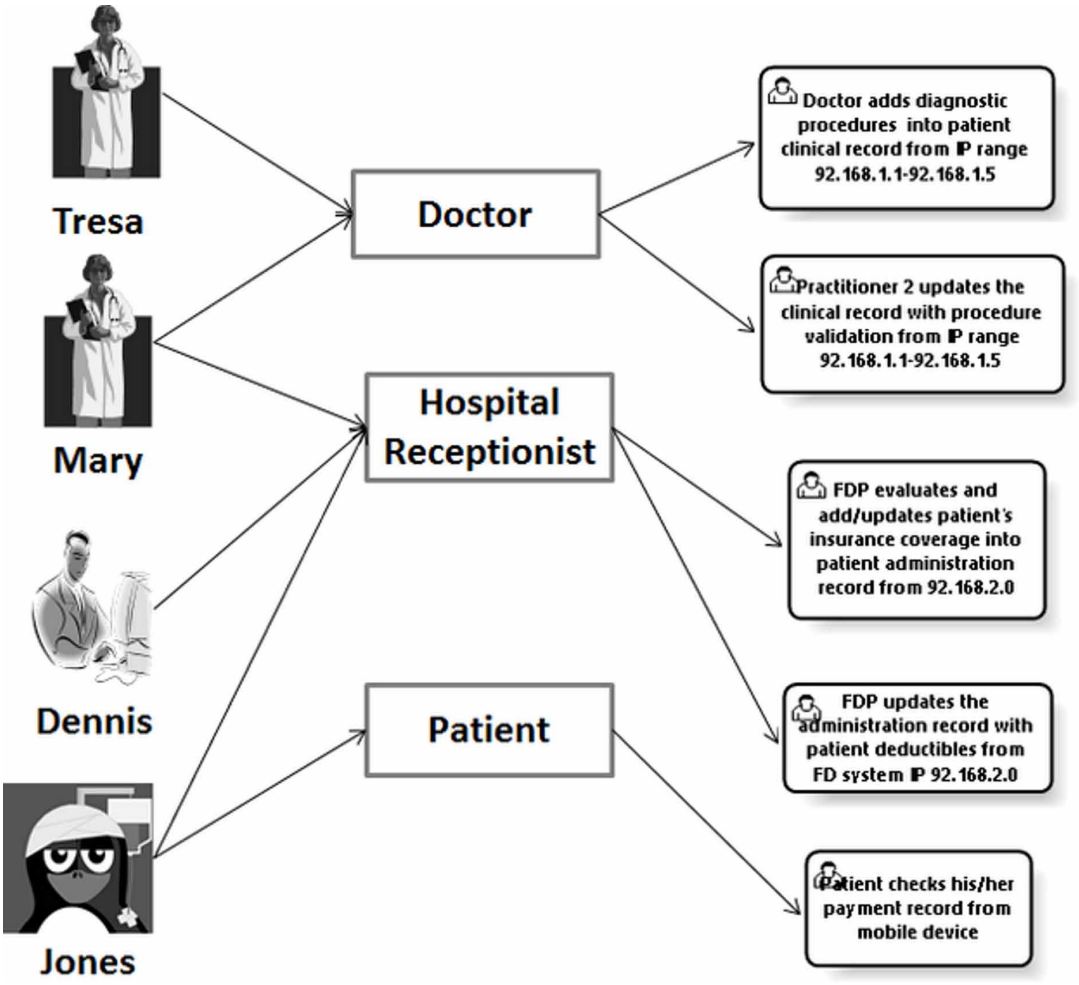
Once the patient has an appointment with the doctor, the workflow in Figure 5will begin. Activities of the workflow given in Figure 5 are discussed in detail. This particular scenario begins when a patient has an appointment with a practitioner and the practitioner recommends a diagnostic procedure.

This workflow starts with (activity A1) identifying the requirements for the diagnostic procedures to be performed (e.g. tests to evaluate a patient's health). Here the doctor will identify the tests and treatments for the patient from ip address 91.91.1.1-91.91.1.5;

Once the diagnostic procedures are identified, the recommended diagnostic procedures will be validated and performed in accordance with the Medicaid/Medicare regulations. This activity (A2) involves conducting the procedure validation to check for appropriate diagnosis in comparison with the Medicaid/Medicare regulations by another practitioner2 from 91.91.1.1-91.91.1.5. After that, if the procedure validation is correct and performed, the workflow will be carried on to Activity 4, otherwise diagnostic procedures will be reviewed and will be recommended to re-assess by the practitioner1 based on the Medicaid/Medicare regulations;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Patient walks in for an appointment



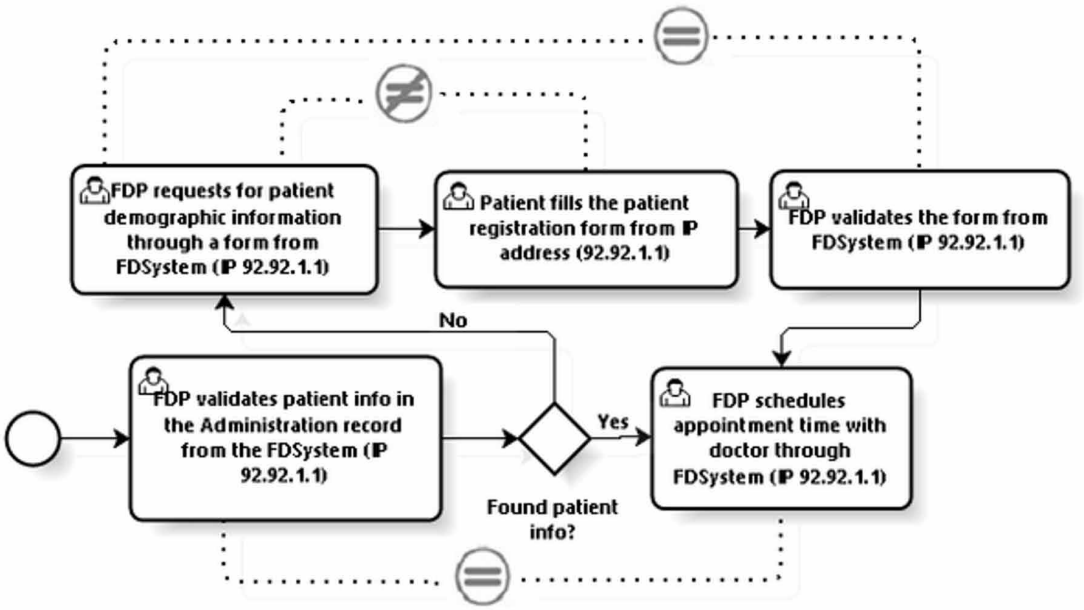
Then the patient's insurance will be evaluated by the front desk person of the hospital staff (e.g. by a receptionist) (A3) from ip address 92.92.1.1. Based on the patient's information, his/her insurance will be evaluated to verify whether his/her current insurance policy would cover that particular test (for example, an eye or a dental exam). If the insurance covers that particular test, workflow will move on to the next activity (A4). If the insurance will not cover the test, the patient may go back and check for a secondary insurance policy which will cover that test;

Activity 4 (A4) is where the financial administrator will calculate the patient's deductibles, once the patient's insurance is validated. This includes the percentage of the total bill that the insurance company would bear (e.g. 85%-15% deductible) vs the amount the patient is liable to pay.

Activity 5 (A5) is where the patient will access his/her payment record. The computed quote is available in the payment record and the patient will access it through a mobile device.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Subject attributes to resource allocation



Attribute based Access Control Policies

The following is a subset of the access control rules in the case study: Some of the ABAC policies for the proposed case study may include:

PPrac: Any user with the position Practitioner/SupervisorPractitioner and an employee of the Organization, and are assigned to the patient, can read/write a Patient record from 91.91.1.1-91.91.1.5

SARE: Target is Subject: Practitioner, Resource: Patient Record, Action: read/write, Environment: 91.91.1.1-91.91.1.5

PN: A Nurse who is working under a practitioner can access patient record from internet network (e.g., IP range: 91.91.1.1-91.91.1.5)

SARE: Target is Subject: Nurse, Resource: patient record, Action: access, Environment: 91.91.1.1-91.91.1.5.

PFDP: Any user with Position: FrontDeskPerson from Department Front Desk with UserID Receptionist can 'access' the 'insurance information' of a patient and thereby 'validate' the Insurance of a patient and add/update it to the payment record, from 92.92.1.1.

SARE: Target is Subject: FrontDeskPerson, Resource: Payment Record, Action: Add/Update, Environment: 92.92.1.1

PPR: Patient Record cannot be accessed by any user with Position: FrontDeskPerson from the IP 92.92.1.1.

If tried to access, email to the Supervisor 'FrontdeskPerson with name X has tried to access Clinical Record X from ip address 92.92.1.1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

SARE: Target is Resource: 'Patient Record', Subject: FrontdeskPerson, action: access, Environment: 92.92.1.1.

PPayR1: Any user with Designation: Patient whose age is greater than 16 can read his/her own Payment Record from internal network or from his/her mobile device at any time.

SARE: Target is Resource: Payment Record, Subject: Patient, Action: read, Environment: Hospital's internal network or from his/her mobile device

PPayR2: Any user with Designation: guardian or parent can read Payment Record of a Patient whose age is less than 16 from hospital's internal network or from his/her mobile device.

SARE: Target is Resource: Payment Record, Subject: Gaurdian/Parent, Action: read, Environment: hospital's internal network or from his/her mobile device.

PIN1: Practitioner/Nurse/SupervisorPractitioner can only read/write patient record from 91.91.1.1-91.91.1.5.

SARE: Target is Subject: Practitioner or Nurse or SupervisorPractitioner, Resource: Patient Record, Action: read/write, Environment: 91.91.1.1-91.91.1.5

PFDD1: Only FDPerson can be logged into 'FDDesktop', and read PatientInsuranceInformation.

SARE: Target is Subject: FDPerson; Resource: Patient Insurance Information, Action: login, Environment: 92.92.1.1;

WSO2

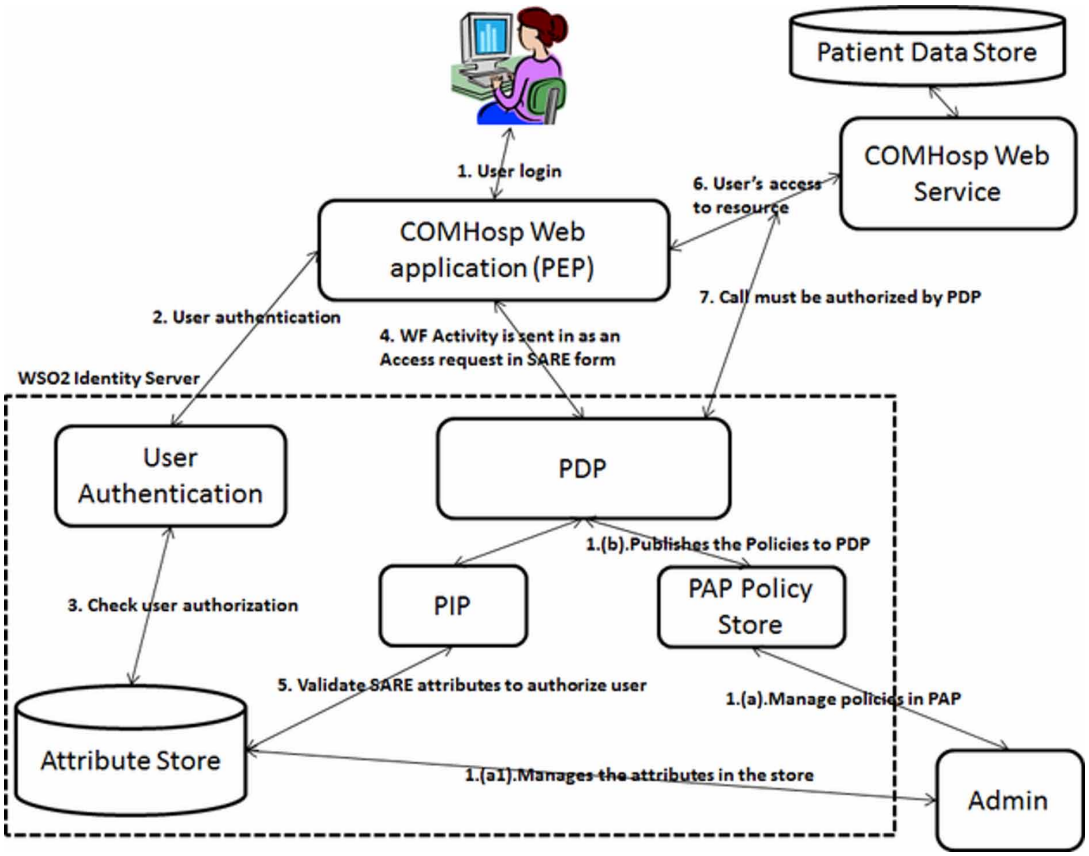
WSO2 identity server is an open source application, which hosts a web service and a web application. It offers a user interface with multiple features where a user can log into the system and create, edit, read, or delete an access control policy. A user can login as an administrator to create and manage access control policies which can be tested in another user interface in which the user can send a request in terms of SARE attributes. For example, if a practitioner tries to access a patient health record, the practitioner's credentials are validated against the access control policy file which is located on a XACML engine in the identity server, and after validation the engine will permit/deny the user's request. A detailed demonstration of the WSO2 identity server with an example is discussed below.

Demonstrating the Case Study using WSO2

The process of validation of policies on workflow activities is shown in the above Figure 7. To validate the model, we have used case study analysis in which we've used Business Process Management Notation (BPMN) to develop the workflow and WSO2 Identity Server to administer the ABAC policies. The steps involved in applying these two for our integrated model are discussed below: ABAC policies written in plain English are entered into the WSO2 Identity server's Policy Administration Point (PAP) through their admin interface, and workflow is developed using BPMN. Once we have policies in WSO2 Identity Server and the workflow, each of the workflow activities is individually identified. Each of the identified workflow activities are extracted from the workflow and the SARE elements are derived from each activity. A request is formulated from the workflow activity's SARE elements, and that request is sent to the WSO2 Identity server's Policy Decision Point. Once the request is sent, WSO2 Identity server will evaluate the request. Based on the policies specification, the server will generate a response containing either 'permit' or 'denied' or 'indeterminate'.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Demonstrating the approach using WSO2



- A 'permit' response will denote that the intended actions by the subject under certain environmental conditions are allowed.
- A 'deny' response will denote that the intended actions by the subject under certain environmental conditions are not allowed.
- An 'indeterminate' response will denote that the activity is not fully covered by the policies defined.

To accommodate complex ABAC policies, more coding is required to modify and accommodate more complex WSO2 features. With some coding into WSO2, it is possible to identify which policies are being enforced on the request (workflow activity).

WSO2 in Action

Here, we'll demonstrate the proposed research methodology using WSO2 Identity server. If we consider activity A1: A Practitioner adds diagnostic procedure into EHR from hospital's internal network, we send this as a request through WSO2 Identity server (in Figure 8). Once the request is evaluated, we can see the result 'permit' in the below Figure 9.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. PPrac Policy in WSO2

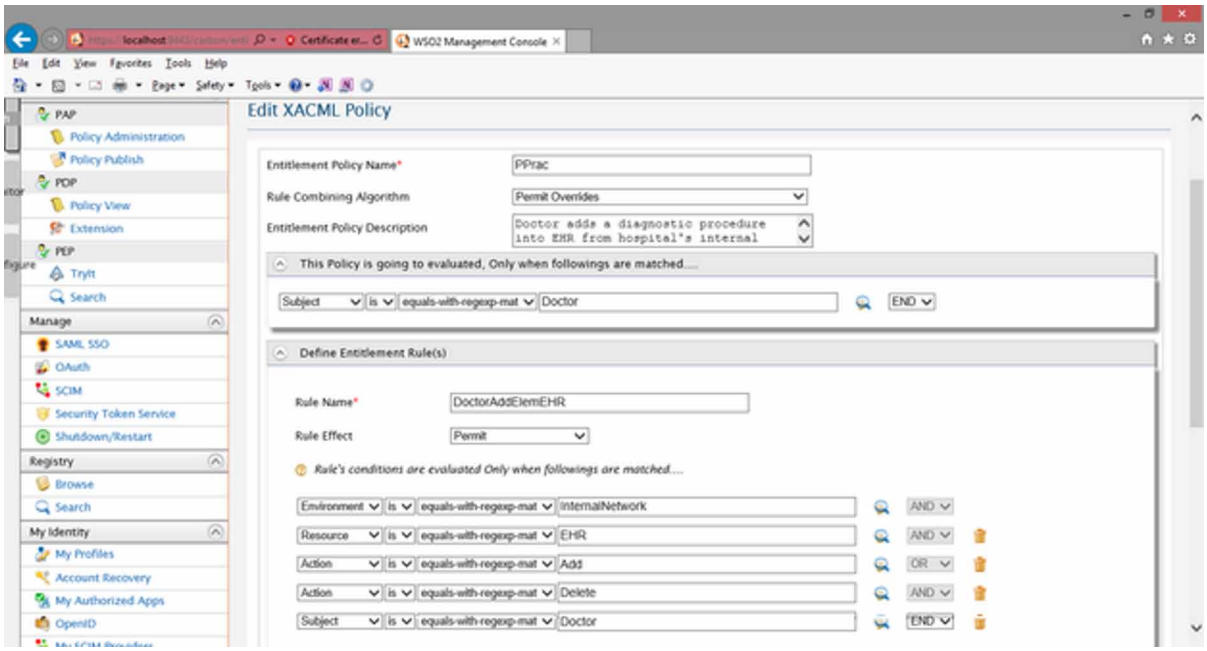
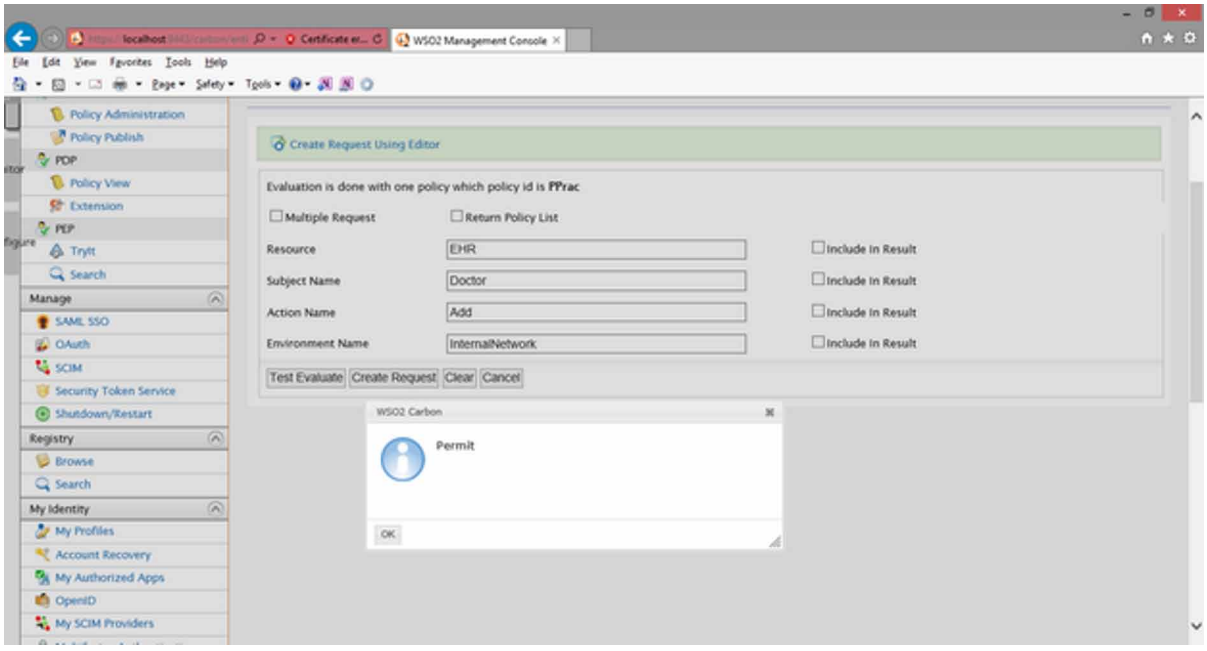


Figure 9. Policy PPrac in action



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Testing Results

All of the available ABAC policies are: PPrac, PPR, PFDP, PPayR1, PPayR2, and PN. The activity analysis helps in understanding that not all of the policies are being used across the activities of the workflow, and can be seen in Table 4. The policies (y) that are not being used at least once include: PN. The policies that are applicable on each activity of the workflow can be found in Table 3 and the policies that are being used (not used) on the workflow activities can be found in Table 4.

GUIDELINES FOR HOSPITALS TO IMPLEMENT THE APPROACH

Purpose

This guide contains the necessary specifications for the integration of workflows and access control policies in a healthcare setting. An integrated methodology will allow efficient and accurate specification and modification of access control policies and workflow activities while reducing the burden on dealing with access control and workflow activities separately. This guide defines how healthcare organizations can define attribute based access control policies; define the implementation of XACML through WSO2; defining and creating a policy administration point; defining the method to extract workflow activities from workflows.

Audience

The intended audiences for this guide are hospital personnel who intend to improve their current access control mechanisms by making environment element their organizational priority. This guide helps in integrating workflow activities and ABAC policies through WSO2 Identity server.

Table 3. Policies applicable on each activity of the workflow

Activity Name	Policy applicable (Yes/ N0/ Partial)?	If applicable, Policy Name
A1	Yes	PPrac, PPR, PIN1
A2	Yes	PPrac, PPR, PIN1
A3	Yes	PFDP, PPayR1, PPayR2, PFDD1
A4	Partial	PPayR1, PPayR2, PFDD1
A5	Partial	PPayR1, PPayR2

Table 4. Policies that are being used (or not used) on the workflow activities

Policy	Policy in use (Y/N)?
PPrac	Y
PPR	Y
PFDP	Y
PN	N
PPayR1	Y
PPayR2	Y
PIN1	Y
PFDD1	Y

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Steps to Implement the Proposed Approach

Assess Current Practice

Some of the elements to assess current practice at a healthcare facility may include: (a) performing an analysis on both workflows and access control policies in the organization; (b) performing an analysis on current type of access control mechanism (c) validating the clinical workflows to check if they are developed based on SARE elements; (d) are access control policies and workflow activities using same SARE elements; (e) does the organization considers environment element as a crucial element.

Setting the Goals

After assessing the current practice, healthcare organization should set clear goals. To implement the proposed methodology, organization is recommended to set and pursue 'specific' 'measurable' 'attainable' 'relevant' and 'time bound' (SMART) goals (Robert, 2005). The healthcare organization is recommended to make the 'environment' element an organizational priority. Some of the goals to be set by the healthcare organization include: (a) identifying and developing SARE elements across the organization, (b) developing workflow activities in terms of SARE elements, (c) developing ABAC policies in terms of SARE elements, (d) developing policy to use same SARE elements across the development teams, (e) making the environment element an organizational priority, (f) providing training and support to the workflow and policy developer team.

Planning the Process

Once the goals are identified, the organizational executive team should develop a plan to implement the methodology and should identify the necessary steps to analyze the process of implementation.

Forming an Expert Team to Identify SARE Elements Which Can Be Used Across the Workflow Development Team and Access Control Developing Team

Healthcare organization should identify an expert team to analyze current working process and identify the required SARE attributes so that the same SARE attributes can be used across the teams that develop workflows and access control policies. The purpose of this team is to define the SARE elements to be used commonly by both workflow and the ABAC policies. This team is recommended to have regular meetings to review and add/modify/delete the existing SARE attributes.

Creating a Policy Statement for SARE Elements Usage Across the Organization

To follow uniform standards of data usage across various electronic platforms, healthcare organization should make a policy statement about SARE elements across the organizational development teams. The organization should also make the policy statement available to the development teams and should work on making the teams to understand the statement.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Training on Developing Workflow Activities in Terms of SARE Elements

Proper training to the workflow developers on how to develop workflow activities in terms of SARE elements should be given. It is an important task to develop workflow activities in terms of SARE elements as these are the elements which would help in integrating the workflow activities with the ABAC policies.

Plan to Implement ABAC through WSO2 Identity Server

As WSO2 identity server is an open source application and accommodates healthcare requirements, WSO2 identity server has been used to implement the approach in this research. As the WSO2 identity server is an open source application and with proper coding knowledge, it can be easily modified based on the organization's requirement. It is recommended by the author to the organization to plan to implement ABAC policies through WSO2 identity server.

Employing an ABAC Policy and WSO2 Identity Server Expert

To implement and develop ABAC policies through WSO2, the developer team should have sufficient knowledge about the concept, which requires having an ABAC policy and WSO2 Identity server expert on team. ABAC policies should be developed from the SARE elements recognized by the expert team. The policy expert can provide required support and guidance to the policy developing team.

Training on ABAC and WSO2 Identity Server

In order to implement ABAC through WSO2 Identity server, the team should have proper training about ABAC concepts, how to implement WSO2, how WSO2 works, how to write ABAC policies in WSO2, how to validate requests through WSO2, and also how to edit the WSO2 identity server.

Summary

This research addresses crucial problem for future healthcare organizations which accommodate environment into workflow activities and access control policies; we have addressed a crucial objective: integrating the workflows and access control policies; and our integrated methodology can identify workflow activities that are not being protected by access control policies. They can further be improved by modifying the workflow activities and access control policies through integrating workflow activities and attribute based access control policies using SARE (Subject, Action, Resource, and environment) elements.

Our main contributions include: (i) proposing 'environment' as a crucial element in a workflow activity along with Subject, Action, and Resource, (ii) a methodology to integrate workflow activities and access control policies, using this methodology to address below questions: (a) Whether there are any workflow activities that are not covered by the access control policies, (b) Whether there are any unused access control policies, (c) How many policies are covering each activity, and (d) Whether the workflow activities and the access control policies need to be modified or refined, (e) Demonstrating the methodology with WSO2 Identity Server; (iv) guidelines for healthcare organizations to implement the approach.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As a part of our future work we have plans to work on: (i) extend the case study to perform in-depth testing and analysis with multiple healthcare scenarios, (ii) modifications to WSO2 identity server are needed to accommodate extended policy sets, (iii) modifications to WSO2 are needed to view the policies that are effecting a workflow activity, (iv) identify any policy related conflict issues, (v) perform in-depth analysis of the integrated methodology.

REFERENCES

Akinyele, J. A., Pagano, M. W., Green, M. D., Lehmann, C. U., Peterson, Z. N. J., & Rubin, A. D. (2011). Securing electronic medical records using attribute-based encryption on mobile devices. *Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices SPSM '11* (pp. 75–86). 10.1145/2046614.2046628

Alhaqbani, B., & Fidge, C. (2007). Access control requirements for processing electronic health records. *Proceedings of the 2007 International Conference on Business Process Management BPM'07* (pp. 371–382). doi:10.1007/978-3-540-78238-4_38

Basin, D., Burri, S. J., & Karjoth, G. (2011). Obstruction-free authorization enforcement: Aligning security with business objectives. *Proceedings of the IEEE Computer Security Foundations Symposium* (pp. 99–113).

Bertino, E., Ferrari, E., & Atluri, V. (1999). The specification and enforcement of authorization constraints in workflow management systems. *ACM Transactions on Information and System Security*, *2*(1), 65–104. doi:10.1145/300830.300837

Campbell, E. M., Sittig, D. F., Ash, J. S., Guappone, K. P., & Dykstra, R. H. (2006). Types of unintended consequences related to computerized provider order entry. *Journal of the American Medical Informatics Association*, *13*(5), 547–556. doi:10.1197/jamia.M2042 PMID:16799128

Chaari, S., Biennier, F., Ben Amar, C., & Favrel, J. (2004). An authorization and access control model for workflow. *Proceedings of the First International Symposium on Control, Communications and Signal Processing*. 10.1109/ISCCSP.2004.1296239

Cole, K. (2002). HIPAA Compliance: Role Based Access Control. *Global Information Assurance Certification Paper - SANS Institute*.

Cole, R., Purao, S., Rossi, M., & Sein, M. K. (2005). Being proactive: where action research meets design research. *Proceedings of the TwentySixth International Conference on Information Systems* (pp. 1–21). Association for Information Systems.

Eekels, J., & Roozenburg, N. F. M. (1991). A methodological comparison of the structures of scientific research and engineering design: Their similarities and differences. *Design Studies*, *12*(4), 197–203. doi:10.1016/0142-694X(91)90031-Q

Ferraiolo, D. F., Sandhu, R., Gavrila, S., Kuhn, D. R., & Chandramouli, R. (2001). Proposed NIST Standard for RBAC. *ACM Transactions on Information and System Security*, *4*(3). doi:10.1145/501978.501980

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ferreira, A., Cruz-Correia, R., Antunes, L., & Chadwick, D. (2007). Access control: How can it improve patients' healthcare? *Studies in Health Technology and Informatics*, *127*, 65–76. PMID:17901600

Giaglis, G. M. (2001). A taxonomy of business process modeling and information systems modeling techniques. *International Journal of Flexible Manufacturing Systems*, *13*(2), 209–228. doi:10.1023/A:1011139719773

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, *28*(1), 75–105. doi:10.2307/249422

HHS.Gov. (n. d.). Health Information Privacy - HIPAA Privacy Rule., U.S. Department of Health & Human Services, HHS.gov (2003).

Hu, V. C., Scarfone, K., & Kuhn, R. (2013). *NIST Special Publication 800-162 DRAFT - FINAL Guide to Attribute Based Access Control (ABAC) Definition and Considerations NIST Special Publication 800-162 DRAFT - FINAL Guide to Attribute Based Access Control (ABAC)*. Definition and Considerations.

Identifier, D. (2005). eXtensible Access Control Markup Language. *Oasis Standard,* February, 23.

Janssen, C. (2014). *Security Breach*. TechoPedia.

Lakkaraju, S. (2013). *Context Aware Mobile Knowledge Management System in Healthcare* (pp. 1–10). MWAIS.

Lakkaraju, S., & Lakkaraju, S. K. (2011). Mobile Device Application in Healthcare. In Cases on Healthcare Information Technology for Patient Care Management (p. 22).

Lakkaraju, S., & Moran, M. (2011). *A Framework to Investigate the Role of Mobile Technology in the Healthcare Organizations*. MWAIS.

Lakkaraju, S., & Xu, D. (2014). Integrated Modeling and Analysis of Attribute based Access Control Policies and Workflows in Healthcare. *Proc. of the 1st International Conference on Trustworthy Systems and Their Applications TSA '14*, Taiwan (p. 8). 10.1109/TSA.2014.15

Le, X. H., Doll, T., Barbosu, M., Luque, A., & Wang, D. (2012). An enhancement of the Role-Based Access Control model to facilitate information access management in context of team collaboration and workflow. *Journal of Biomedical Informatics*, *45*(6), 1084–1107. doi:10.1016/j.jbi.2012.06.001 PMID:22732236

Leyva, C. (2014). Hitech Act Summary. *HIPAA Survival Guide*. Retrieved from http://www.hipaasurvivalguide.com/hitech-act-summary.php

Ligatti, J., Bauer, L., & Walker, D. (2005). Edit automata: Enforcement mechanisms for run-time security policies. *International Journal of Information Security*, *4*(1-2), 2–16. doi:10.100710207-004-0046-8

Lowe, G. (1996). Some new attacks upon security protocols. *Proceedings of the 9th IEEE Computer Security Foundations Workshop* (pp. 162–169). 10.1109/CSFW.1996.503701

Lu, Y., Zhang, L., & Sun, J. (2009). Task-activity based access control for process collaboration environments. *Computers in Industry*, *60*(6), 403–415. doi:10.1016/j.compind.2009.02.009

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. doi:10.1016/0167-9236(94)00041-2

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. doi:10.2753/MIS0742-1222240302

Robert, B. (2005). Use S.M.A.R.T Goals to Launch Management by Objectives Plan. *TechRepublic*. Retrieved from http://www.techrepublic.com/article/use-smart-goals-to-launch-management-by-objectives-plan/

Roscoe, A. W. (2005). The Theory and Practice of Concurrency. *Theory into Practice*, *216*.

Russello, G., Dong, C. D. C., & Dulay, N. (2008). A Workflow-Based Access Control Framework for e-Health Applications. *Proceedings of the 22nd International Conference on Advanced Information Networking and Applications - Workshops (AINA Workshops '08)*. 10.1109/WAINA.2008.131

Sandhu, R. (1988). *Transaction control expressions for separation of duties*. Fourth Aerospace Computer Security Applications. doi:10.1109/ACSAC.1988.113349

Schneider, F. B. (2003). Enforceable security policies. *Proceedings of the Foundations of Intrusion Tolerant Systems '03*.

Siriwardena, P. (2013). Understanding HIPAA. *Facilelogin*. Retrieved from http://blog.facilelogin.com/2013/02/understanding-hipaa.html

Sittig, D. F., Krall, M., Kaalaas-Sittig, J., & Ash, J. S. (2005). Emotional Aspects of Computer-based Provider Order Entry: A Qualitative Study. *Journal of the American Medical Informatics Association*, *12*(5), 561–567. doi:10.1197/jamia.M1711 PMID:15905478

Spear, N., Malladi, S., & Lakkaraju, S. (2012). Analyzing Workflows in Business Processes for Obstructions Due to Authorization Policies. *Proceedings of the 45th Hawaii International Conference on System Sciences* (pp. 5340–5349). 10.1109/HICSS.2012.112

Tan, K., Crampton, J., & Gunter, C. A. (2004). The consistency of task-based authorization constraints in workflow. *Proceedings of the 17th IEEE Computer Security Foundations Workshop* '04.

United States Department of Defense DoD. (1985). Trusted Computer System Evaluation Criteria

Van Der Aalst, W.M.P. (1998). The application of petri nets to workflow management. *Journal of Circuits, Systems, and Computers*.

van der Aalst, W. M. P. (2004). Business process management: A personal view. *Business Process Management Journal*, *10*(2), 135–139. doi:10.1108/bpmj.2004.15710baa.001

Welch, J. (2014). Improving Clinical Workflow: An Example from the Emergency Department. *Health Catalyst*. Retrieved from https://www.healthcatalyst.com/improve-clinical-workflow-emergency-department

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Wolter, C., Schaad, A., & Meinel, C. (2008). Task-based entailment constraints for basic workflow patterns. *Proceedings of the 13th ACM symposium on Access control models and technologies - SACMAT '08* (pp. 51–60). ACM Press. 10.1145/1377836.1377844

XiangPeng. Z., Cerone, A., & Krishnan, P. (2006). Verifying BPEL Workflows Under Authorization Constraints. *Proceedings of the 4th International Conference Business Process Management BPM* (pp. 439–444).

Xu, D., & Zhang, Y. (2014). Specification and Analysis of Attribute-Based Access Control Policies: An Overview. *Proceedings of the International Workshop on Information Assurance, in Conjunction with the 8th International Conference on Software Security and Reliability (SERE'14).* 10.1109/SERE-C.2014.21

Yuan, E., & Tong, J. (2005). Attributed based access control (ABAC) for Web services. *Proceedings of the IEEE International Conference on Web Services ICWS '05*. 10.1109/ICWS.2005.25

Zhang, G., & Liu, J. (2011). A Model of Workflow-oriented Attributed Based Access Control. *International Journal of Computer Network and Information Security*, *1*(February), 47–53. doi:10.5815/ijcnis.2011.01.07

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 11(1); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 1-20; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 70
Big Data and IoT–Allied Challenges Associated With Healthcare Applications in Smart and Automated Systems

Mamata Rath
C.V. Raman College of Engineering, India

ABSTRACT

Most of the current smart applications are developed with basis on intelligent computing, most of which are implemented using big data analytics and much other advanced technology. With emerging technology, industrial and instructive improvements are causing greater changes in the lifestyle of people in smart cities and there is more chance of various health problems in urban areas. The way of life in metro urban areas with an expansive volume of people is similarly influenced by different applications and administration frameworks. In this way, the majority of the urban communities are transforming into smart urban areas by receiving mechanized frameworks in every conceivable segment. Therefore, there is more health-related issues and health hazard issues can be identified in urban areas. This article carries out a comprehensive survey of health care issues and improved solutions in automated systems using Big Data Analytics in smart cities integrated with IoT.

1. INTRODUCTION

Implementation of smart concepts like of smart homes, smart cities, and smart everything emerged the Internet of Things (IoT) as an area of inconceivable impact, prospective, and growth. The wide-scale dispersion of the Internet has been the main thrust for this developing pattern, to be specific the utilization of such worldwide communication foundation for empowering machines and brilliant articles to impart, coordinate, and take choices on genuine word circumstances. Nowadays, the evolution of the world of the Internet of Thing is promising the explosion of a number of devices connected to the

DOI: 10.4018/978-1-7998-1204-3.ch070

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Internet. According to Cisco analysts (Scarfo et al., 2018) there were more than 25 billion devices in 2015 connected and with a projection of more than 50 billion devices connected by 2020. Also, the new business paradigms that the Internet of Things technologies enable are producing a super-fast increase of machine-to-machine communications. This is a real market breakthrough moment that opens up a lot opportunities for enterprises and, generally speaking, for the whole society. Inherently, it increases dramatically the security problems, which could frustrate a sizeable part of Internet of Things' potential benefits that McKinsey valuates are very high. Indeed, a recent survey by HP reports that the 70% of devices contain vulnerabilities. The aim of this section is to provide an overview of current trends about cyber security concerns and a glimpse of what the future of the Internet of Things will bring.

Big data is used in several research areas related to healthcare, area-based services, satellite data usage, online advertising, and retail marketing. In the coming years, the Internet of Things (IoT) will increase the measure of data on the planet, and an exponential rise in big data will be seen. There have been a few review papers on big data in general and in diverse specialized fields, e.g., science, healthcare, geography, and the Internet. Unlike previous reviews in the literature, this paper takes a gander at recent advances in big data from a different view and with different arrangements, i.e., data types, storage models, investigation models, protection, data security, and applications.

2. BIG DATA AND IOT INCORPORATED APPLICATIONS

Internet of Things (IoT) is a transformational technology for electronic security frameworks. In numerous regards, business and private security items were forerunners to IoT and keep on sharing numerous critical qualities of the class. All things considered, the fast drop in cost for IoT gadgets and the billions of new IoT gadgets that are relied upon to be introduced in the following 5– 10 years make it a power to be battled with as far as how we consider electronic security frameworks and the market all in all. IoT additionally enhances the industry tasks in light of the fact that the most every now and again referred to shopper profit of IoT in the house is for security applications. By and large, the ascent of IoT increment value for security purchasers on the grounds that the financial aspects of the IoT industry will apply descending cost weight (Qiu et al., 2015) on security segments, even as they turn out to be more competent and give more highlights.

Regardless of how hard directions and the intricacy of human science makes improving in medication, the pace of advancement in social insurance is incredibly quick. Declarations and news astonish us consistently, regardless of whether we grew up perusing and watching sci-fi. Organizations have effectively printed out liver and kidney tissues; tweaked prosthetics and even medications endorsed by the FDA. IBM Watson's computerized reasoning gathers gigantic measures of data and plans the best treatment alternative for patients by checking all important medicinal examinations, and profound learning calculations will do much more. Advanced tattoos can gauge the significant health care parameters and essential signs to tell us when there's brief comment care of through our cell phones. Increased reality gadgets, for example, Hololens from Microsoft can extend advanced scenes on what we see and get ready specialists for troublesome strategies. Medicinal services development simply doesn't quit shocking people. In 2013, there were bits of gossip about new businesses reforming adherence control by embeddings microchips into sedate containers to give doctors a chance to check whether and when patients take them. We have been sitting tight for gadgets that permit holographic information input yet just oversimplified, toy-like contraptions have turned out to be accessible. Without extensive organiza-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tions putting resources into them, they will have far to go before standard selection. An extraordinary case is L'Oréal, the beauty care products organization putting into a wearable sensor that fills the wearer in regarding whether their introduction to the Sun achieves risky levels.

As of late, the Smart City idea has turned out to be prevalent for its guarantee to enhance the personal satisfaction of urban residents. The idea includes different controls, for example, Smart social insurance, Smart transportation, and Smart people group. Most administrations in Smart Cities, particularly in the Smart medicinal services space, require the constant sharing, preparing and dissecting of Big Healthcare Data for astute basic leadership. In this manner, a solid remote and versatile correspondence foundation is important to associate and access Smart social insurance administrations, individuals and sensors consistently anyplace whenever. In this situation, versatile distributed computing (MCC) can assume a crucial part by offloading Big Healthcare Data related errands, for example, sharing, handling and examination, from portable applications to cloud assets, guaranteeing nature of administration (QoS) requests of end-clients. Such asset movement, which is likewise named Virtual Machine (VM) relocation, is exceptionally compelling in the Smart social insurance situation in Smart Cities.

The immense development of wireless communication has acquired an improved wireless communication network, for the present worldwide communication showcase. These sorts of network are equipped for working with no settled establishment not at all like cell framework. As a result of their self-sufficient, conveyed and multi bounce handing-off potential, mobile ad-hoc networks otherwise known as MANET (Rath et al., 2017) and other wireless networks, are on enormous interest for efficient recuperation, vehicle following and frontline administration. The network topology changes powerfully because of unusual hub developments. This makes a genuine level of challenges to keep up network availability with neighbouring substances. By the day's end, there is no assurance of solid information conveyance with satisfactory Quality of Services (Rath et al., 2018). Steering conventions assume an essential part in setting up end to end way, as hubs are not generally straightforwardly reachable to each other. The power-hungry elements over such data transfer capacity restricted network request hearty and productive directing methodologies for network life span. In this way, directing instrument should be outlined sensibly to streamline network operation. Routinely, there are two primary sorts of directing systems which are table driven and on request premise. Table based methodologies unnecessarily surges the network even without the nearness of real activity and in this way creates tremendous overhead. The wireless communication network having deficient data transmission can't generally manage the cost of the advantage of such extra payload. Despite what might be expected, any on request convention, for example, wireless ongoing begins involving the network resources (Rath et al., 2017) at whatever point there is a need from the clients to pass on any data to their associates. Wireless constant convention keeps quiet until the point when such interest originates from any network substance. Once the presence of previously mentioned undertaking is detected, wireless constant begins route disclosure methodology by communicating a route request parcel. This route request message goes through the network and reaches to real goal and structures a turnaround route table. The beneficiary at that point reacts back by producing a route reply parcel which is conveyed by the turn around route made before. This procedure makes the forward route for information parcels. Here, the route revelation plot prompts delay and furthermore some moderate hubs get over-burden because of doing the information dealing with much of the time. Along these lines, conventional directing strategies can't be used for versatile specially appointed networks straight away on account of their characteristic impediments. In this paper, we focused on tweaking the essential wireless steering convention to provide food the forceful QoS (Quality of Service) (Rath et al., 2017) request of ongoing network benefits in wireless communication. Constant information transmission

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in remote correspondence arrange is a testing issue because of the self-coordinated nature of cell phones which are associated with each other by remote connections. Hubs habitually changes their position and area by shaping a system with no base or any foundation. Cautious contemplations must be made amid planning the conventions for ongoing information transmission which are profoundly touchy. Regular difficulties looked by continuous information amid transmission are vitality use, versatility of the system, improved utilization of data transfer capacity, asset reservation, lessening possibility of connection disappointment, security, deal with QoS, powerfully changing topology and so on.

3. FOCUS ON QUALITY OF SERVICE IN HEALTH CARE SYSTEMS

These days, electronic Health mind services have been accepting increasingly consideration, because of the generously expanding number of superior processing gadgets interconnected with restorative sensors indicative frameworks through heterogeneous network advances. Portable Health mind is a sub-section of the eHealth field, expecting to utilize cell phones, wearable or implicit sensors with Body Area Networks and the vastest size of wireless correspondence advances to upgrade conventional restorative services. Propelled portable advancements assume an inexorably vital part in health care frameworks particularly in situations of tele-conference telediagnosis and versatile patient checking (Li et al., 2007). Normally, portable health care services require strict, therapeutic level QoS and QoE arrangement. The constant utilize instances of portable health care services, for example, remote versatile patient checking, telecare and remotely guided surgical mediation require significantly higher certifications (e.g., little deferral and jitter, quick reaction time, and low bundle misfortune). Numerous portable health care situations depend on wearable body key (e.g., ECG, heart rate screen, and ultrasound) or inherent sensors of advanced mobile phones (e.g., high-determination camera and whirligig). The very shifting qualities of various versatile health care applications in methods for the required network assets, QoS/QoE prerequisites conjure elaboration of cutting edge network administration structures. These actualities persuaded us to outline and actualize a network-helped wireless access network determination system for mHealth (mobile health care) services, which can choose the fitting accessible access network utilizing a multi-criteria choice motor. There is some exploration commitment in which the portable health care framework depends on the Distributed Decision Engine (DDE) and the Network Information Service (NIS) giving both static and dynamic parameters of the accessible human services correspondence networks. There are some methodologies in which the arrangement is executed for Wi-Fi networks. E-health mind frameworks additionally gives a proficient domain to therapeutic quality information transmission in different portable health care situations.

3.1. Big Data in Wireless Body Sensor Networks

Wireless sensor innovations, particularly wireless body area network and sensor networks have gone past the fine-grained constant checking stages and wound up plainly one of the empowering advancements that give numerous effective applications in therapeutic and non-medicinal fields (Zuhra et al., 2017). The physiological observing frameworks are produced to screen the human medical problems and are dependable to course the detected information (physical or imperative information for example, glucose, Electromyography (EMG), Electrocardiogram (ECG), Electro-encephalograph (EEG), temperature, and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

so on.) from biosensors hubs to the medicinal or non-restorative server for encourage examination. Many steering and information spread conventions have been particularly intended for WBSNs. Directing conventions in WBSNs may contrast contingent upon the application and network design.

With the taking off enthusiasm for the Internet of Things (IoT), some human services suppliers are encouraging remote care conveyance using wearable gadgets. These gadgets are utilized for constant spilling of individual therapeutic information (e.g., vitals, solutions, sensitivities, and so on.) into social insurance information frameworks for the reasons for health care checking and proficient conclusion. In any case, a test from the viewpoint of the doctors is the failure to dependably figure out which information has a place with who progressively. This test exudes from the way that medicinal services offices have various clients who claim different gadgets (Lomotey et al., 2017); consequently, making an information source heterogeneity and complexities for the gushing procedure. As a feature of many research, this issue has been streamlined by proposing some wearable IoT information gushing design that offers traceability of information courses from the beginning source to the health care information framework.

3.2. Design of Various Smart IoT Framework in Health Care

The upsides of Internet of Things (IoT) innovation in healthcare division and medical applications have colossally influenced the testing health care issues. Identity Management is a vital strategy in the Internet of Things (IoT), which has been utilized as a part of in an extremely encouraging zone of medical science. In this Identity Management strategy, everything, any client or any gadget, is basically doled out a novel identity, and all data with respect to the gadget can be assembled through its identity.

Numerous Real time system observing are finished utilizing IoT idea with a specific end goal to remain cautioned any sort of action at the gadget condition so vital move can be made instantly subsequent to getting warning about the occasion (Sun et al., 2012). An issue of IoT in mining region is likewise vital issue which is focused here, significant operations of which are completed in mining regions, winding up as tailings. To keep mine following dam disappointments and significant mishaps because of it in which there are ghastly harm and demise happening at a moderately higher rate.

Internet of Things (IoT) offers a consistent stage to interface people and articles to each other for enhancing and making our lives less demanding. This vision conveys us from process based concentrated plans to a more disseminated condition offering an immense measure of uses, for example, keen wearable, shrewd home, splendid movability, and wise urban areas (Farahani et al., 2018). A few articles examine pertinence of IoT in human services and medication by showing a widely inclusive engineering of IoT eHealth organic group. Human services are ending up progressively difficult to oversee because of lacking and less powerful social insurance services to meet the expanding requests of rising developing populace with endless infections. In a few presentations the examiners suggest that e-social insurance framework requires a change from the center driven treatment to persistent driven medicinal services where every operator, for example, facility, patient, and services are consistently associated with each other. This patient-driven IoT eHealth organic group needs a multi-layer engineering: (1) gadget, (2) haze processing and (3) cloud to engage treatment of complex data to the extent its assortment, speed, and lethargy. This fog driven IoT design is trailed by different case cases of services and applications that are actualized on those layers. Those representations keep running from versatile health care, helped living, e-medication, inserts, early forewarning frameworks, to populace checking in shrewd urban areas. The difficulties of IoT eHealth are data organization, versatility, directions, interoperability, gadget network human interfaces, security, and protection.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.3. Confidentiality and Security in IoT Based Health Care Systems

Today, telemedicine has a great reputation because of its capacity to provide quality healthcare services to remote locations. To achieve its purposes, telemedicine utilizes a number of wireless technologies as well as the Internet of Things (IoT) (Hossain et al., 2017). The IoT is redefining the capacity of telemedicine in terms of improved and seamless healthcare services. In this regard, the set of features of telemedicine by proposing a model for an IoT-based health prescription assistant (HPA) have been discussed in, which helps each patient to follow the doctor's recommendations properly. This paper also designs a security system that ensures user authentication and protected access to resources and services. The security system authenticates a user based on the OpenID standard. An access control mechanism is implemented to prevent unauthorized access to medical devices. Once the authentication is successful, the user is issued an authorization ticket, which this paper calls a security access token (SAT). The SAT contains a set of privileges that grants the user access to medical IoT devices and their services and/or resources. The SAT is cryptographically protected to guard against forgery. A medical IoT device verifies the SAT prior to serving a request, and thus, ensures protected access. A prototype of the proposed system has been implemented to experimentally analyze and compare the resource efficiency of different SAT verification approaches in terms of a number of performance metrics, including computation and communication overhead.

4. HEALTH CARE IN IOT NETWORK AND SMART COMMUNITY

Wireless body sensor networks are remote systems that are intended to permit correspondence among sensor hubs that are joined to a human body to screen the body's key parameters and condition. The plan and advancement of such WBSN frameworks for health care checking have gotten a lot of consideration as of late (Harbouche, 2017), in inquire about examinations and in industry. This consideration is for the most part propelled by expensive medicinal services and by late advances in the improvement of smaller than normal health care observing gadgets and in addition developing advances, for example, the Internet of Things (IoT), which add to the primary difficulties of 5G. The presence of an unequivocal way to deal with address the required programming outline and check ought to be exceptionally valuable for the development and upkeep of such frameworks. This paper introduces a preventive human services framework that has an adaptable plan. The proposed framework depends on an engineering that has heterogeneous hubs and gives both day by day persistent checking and additionally particular controls. We characterized a model to depict the WBSN's worldwide conduct. A critical part of this work is that propose a model-driven building (MDE) way to deal with address the determination of every hub's conduct in the WBSN from the WBSN worldwide conduct. This approach enables engineers to acquire a framework plan from the worldwide detail of its necessity. To guarantee the conformance of this outline to its determination, the inferred practices ought to be approved and checked before their arrangement. Truth be told, formal strategies are effective apparatuses for programming specialists to check the consistent accuracy of simultaneous programming at various levels of its life cycle. Display checking is a standout amongst the most capable formal techniques for confirming the consistent accuracy of such simultaneous frameworks. In this work, we make utilization of a model checking approach that depends on a model change to approve the consequently determined conduct of a WBSN for health

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

care observing. This model-driven approach will check whether the determined framework carries on accurately as indicated by its worldwide detail, while the goal is to build the framework's execution and QoS. This approach enables the designer to reason about a model of the worldwide framework instead of about the framework itself.

4.1. Integration of Wireless Network and IoT

The steady progression and change of low-control remote correspondence advancements is engaging the ascent of various new applications (Patel et al., 2017) in the field of the Internet of Things (IoT). One of the essential areas of research inside the Smart City setting is keen prosperity, which attracts novel IoT exercises to improve both quality and access to social protection and astute organizations all around. In CUIDATS, an IoT mutt checking system for restorative administrations circumstances has been presented which facilitates RFID and WSN advancements in a single stage giving territory, status, and following of patients and assets. After its execution endorsement in a suitable testbed, CUIDATS has been passed on and surveyed with an abnormal state of accomplishment in a veritable recuperating focus.

The continuous development and improvement of low-power wireless communication technologies is enabling the emergence of many new applications in the field of the Internet of Things (IoT). One of the main areas of research within the Smart City context is smart health, which engages novel IoT initiatives to improve both quality and access to health care and smart services in general. In CUIDATS, (Adame et al., 2018) an IoT hybrid monitoring system for health care environments has been presented which integrates RFID and WSN technologies in a single platform providing location, status, and tracking of patients and assets. After its performance validation in a suitable testbed, CUIDATS has been deployed and evaluated with a high degree of success in a real hospital.

Ceaseless checking of unending patients enhances their personal satisfaction and decreases the financial expenses of the sterile framework. Be that as it may, with a specific end goal to guarantee a decent observing, high data transmission and low deferral are required (Patel et al., 2017). The 5G innovation offers higher transmission capacity, bring down deferrals and bundles misfortune than past advancements. This paper exhibits an engineering for savvy eHealth observing of unending patients. The engineering components incorporate wearable gadgets, to gather measures from the body, and a cell phone at the patient side so as to process the information got from the wearable devices (Lloret et al., 2017). Additionally. a database is required with a canny framework ready to send an alert when it distinguishes that it is going on something atypical. The astute framework utilizes machine taking in Big Data taken from various clinics and the information got from the patient to analyze and create alerts. Investigation tests have been done to reproduce the movement from numerous clients to the database keeping in mind the end goal to assess the appropriateness of 5G in engineering (Lloret et al., 2017). At the point when there are couple of clients (under 200 clients), we don't discover enormous contrasts of round excursion time in the vicinity of 4G and 5G, however when there are more clients, similar to 1000 clients, it increments significantly achieving 4 times more in 4G The Packet Loss is practically invalid in 4G until 300 clients, while in 5G it is conceivable to keep it invalid until 700 clients . The Outcomes call attention to that keeping in mind the end goal to have high number of patients ceaselessly checked, it is important to utilize the 5G organize in light of the fact that it offers low postponements and ensures the accessibility of data transmission for all clients.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4.2. Ethical Issues in Big Data Management in Healthcare

Prosperity of growth patients and survivors is a test around the world, thinking about the regularly incessant nature of the malady. Today, countless, items and administrations are accessible that mean to give systems to confront the test of prosperity in disease patients; by the by the proposed arrangements are frequently non-economical, exorbitant, inaccessible to those in require, and less generally welcomed by patients. These difficulties were considered in planning FORECAST (Kyriazakos et al., 2017), a cloud-based customized shrewd virtual training stage for enhancing the prosperity of malignancy patients. Customized training for disease patients concentrates on physical, mental, and enthusiastic concerns, which FORECAST can distinguish. Growth patients can benefit from training that tends to their passionate issues, encourages them concentrate on their objectives, and backings them in adapting to their illness related stressors. Customized training in FORECAST offers bolster, consolation, inspiration, certainty, and seek and is an important instrument after the prosperity of a patient (Kyriazakos et al., 2017).

The ability to incorporate and examine gigantic measures of data is critical to the achievement of associations, including those that include worldwide health care. As nations turn out to be exceedingly interconnected, expanding the hazard for pandemics and flare-ups, the interest for big data is probably going to increment. This requires a worldwide health care workforce that is prepared in the compelling utilization of big data. To evaluate usage of big data preparing in worldwide health care, we led a pilot review of individuals from the Consortium of Universities of Global Health. The greater part the respondents did not have a big data preparing program at their institution. Additionally, the majority concurred that big data preparing projects will enhance worldwide health care expectations, among other positive results. Given the watched hole and benefits, worldwide health care teachers may consider putting resources into big data preparing for understudies looking for a profession in worldwide health (Olayinka et al., 2017).

With the EU wide-ranging Data security instruction going into oblige in 2018, the stage is set for worldwide verbal come across on Big Data sharing in prosperity. Taking into consideration the way that prosperity information (and especially innate information) are seen as "touchy", is there a way to deal with structure the verbal encounter on the obstructions, and danger advantage extent that moves a long way from the conventional points of interest and burdens of potential insurance and isolation risks (Knoppers et al., 2017). Potential partition has been tended to in sanctioning and the altering of security rights against the potential advantages of information sharing in concentrated science is provoking a more proportionate approach. We suggest that a basic stimulus that will also move this practical talk from its customary structures is start the "right to science" as appreciated in general law. Of note, the Global Alliance for Genomics and Health has developed a Framework for Responsible Sharing of Genomics and Health-Related Data in perspective of that human right. So likewise, late standards from the Council for International Organizations of Medical Sciences and the Organization for Economic Co-operation and Development, and what's more from the American College of Medical Genetics and Genomics contain courses of action that propel information sharing showing that information concentrated science may a little bit at a time come to be set up on a more aggregate ethos.

In the period of "big data", late improvements in the territory of data and correspondence advances (ICT) are facilitating associations to enhance and develop. These innovative improvements and wide adjustment of ubiquitous figuring empower various opportunities for government and organizations to reexamine medicinal services prospects. Along these lines, big data and brilliant medicinal services frameworks are autonomously pulling in broad consideration from both scholarly community and in-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

dustry (Pramanik et al., 2017). The blend of both big data and brilliant frameworks can expedite the possibilities of the medicinal services industry. Notwithstanding, an intensive investigation of big data and shrewd frameworks together in the social insurance setting is as yet missing from the current literature. The key commitments of this article incorporate a sorted assessment of different big data and shrewd framework advances and a critical examination of the best in class propelled medicinal services frameworks. We depict the three-dimensional structure of a change in outlook. We additionally remove three wide specialized branches (3T) adding to the advancement of human services frameworks. All the more particularly, we propose a big data empowered savvy medicinal services framework structure (BSHSF) that offers hypothetical portrayals of an intra and bury authoritative plan of action in the social insurance setting. Pramanik et al. (2017) likewise specify a few illustrations announced in the literature, and afterward we add to pinpointing the potential opportunities and difficulties of applying BSHSF to human services business situations. We likewise make five proposals for viably applying `BSHSF to the social insurance industry. To the best of our insight, this is the first inside and out examination about cutting edge big data and shrewd social insurance frameworks in parallel. The administrative ramifications of this article is that associations can utilize the discoveries of our critical examination to fortify their vital plan of shrewd frameworks and big data in the human services setting, and thus better use them for feasible hierarchical creation.

A major information examination engaged change show in light of preparing based view is made, which reveals the causal associations among huge information investigation capacities, IT-enabled change sharpens, advantage estimations, and business values (Wang et al., 2017) . This model was then attempted in human administrations setting. By analyzing huge information use cases, we attempted to perceive how huge information examination capacities change legitimate practices, along these lines making potential advantages. Notwithstanding sensibly describing four major information examination capacities, the model offers a fundamental point of view of enormous information investigation. Three basic approach to-regard chains were recognized for social protection relationship by applying the model, which gives helpful encounters to boss.

4.3. Features of Smart and Connected Communities (SCC) in Health Care

Information experts are tested to plan health care data innovation (IT) answers for complex issues, for example, health care disparities, yet are accomplishing blended outcomes in exhibiting an immediate effect on health care results. This introduction of aggregate insight and the relating terms of shrewd health care, learning biological community, improved health care disparities informatics capacities, learning trade, big-data, and situational mindfulness are methods for showing the perplexing difficulties informatics experts look in attempting to model, measure, and deal with a canny and savvy frameworks reaction to health care disparities (Carney et al., 2017). A critical piece in our comprehension of aggregate insight for open and populace health care rests in our comprehension of open and populace health care as a living and developing system of people, associations, and assets. This exchange speaks to a stage in propelling the discussion of what a shrewd reaction to health care disparities should speak to and how informatics can drive the plan of wise frameworks to help with killing health care disparities and accomplishing health care equity.

Development of setting mindful advancements and IoT gadgets mirror that the quality of a human life has turned out to be a standout amongst the most basic angles in Smart Cities. With this objective health care monitoring of elderly and incapacitated individuals have a lot of consideration and center in

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the examination. The social insurance frameworks depend on the parts in charge of setting detecting, handling, stockpiling and derivation, and reaction. Keeping in mind the end goal to make the interoperability among the different human services frameworks, a regular standard is required so as to consistently get to the setting mindful social insurance data (Hussain et al., 2015) getting through a key foundation. An individual-driven detecting system for the human services of elderly and impaired individuals has been proposed, where such stage is planned to monitor strength of the elderly and debilitated individual and furnish them with an administration situated crisis reaction in the event of irregular health care condition. We concentrate on three viewpoints: (a) setting control from the cell phone in individuals driven condition; (b) crisis reaction utilizing setting base data; and (c) displaying portable setting sources as administrations. The most particular element of current work is that medicinal assets are productively used to give them ongoing restorative administrations if there should arise an occurrence of crisis at the same time expanding interpersonal organization of the elderly individuals. The framework usage demonstrates that the proposed individuals driven detecting framework is proficient and financially savvy in health care and crisis mind.

The immediate combination of sensors within the wrap to frame a brilliant dressing has been proposed as conceivable methods for handling a portion of the present difficulties in the treatment of incessant injuries. The improvement of an "associated health care" demonstrate (Finlay et al., 2016) through which the injury is remotely monitored within the community and surveyed by a clinician or independently by a proper brilliant application algorithm has many points of interest, however it is especially a "perfect." Although there are numerous indicative measures and detecting systems that can be interpreted from the lab seat to the center, the resulting adjustment for persistent/intermittent monitoring that would be required for outpatient monitoring faces various obstacles. The transition of the detecting innovation from the lab seat to the facility or to a patient's home requires the joining of the organic, compound, and materials components with present day correspondence. The coming of the advanced cell and tablet guarantees to upset human services offering customized pharmaceutical and more prominent patient mindfulness and administration of their condition. This part investigates the mechanical prerequisites important to process the flag from the sensor that is in contact with the injury, the mechanisms through which the flag can be transmitted and the more extensive ramifications of developing eHealth and mHealth models.

4.4. Big Data Analytic Method in Health Care Using IoT

The term biosensors includes gadgets that can possibly evaluate physiological, immunological and behavioral reactions of domesticated animals and various creature species. Novel biosensing approachs offer exceedingly particular monitoring gadgets for the particular estimation of individual and numerous parameters covering a creature's physiology and additionally monitoring of a creature's situation (Neethirajan et al., 2017). These gadgets are not just very particular and sensitive for the parameters being dissected, yet they are additionally solid and simple to utilize, and can quicken the monitoring procedure. Novel biosensors in domesticated animal administration give noteworthy benefits and applications in ailment location and confinement, health care monitoring and recognition of conceptive cycles, and in addition monitoring physiological prosperity of the creature by means of examination of the creature's condition. With the improvement of incorporated frameworks and the Internet of Things, the consistently monitoring gadgets are required to end up noticeably reasonable. The data created from coordinated domesticated animals monitoring is foreseen to help ranchers and the farming business to enhance creature productivity later on. The data is required to lessen the effect of the domesticated

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

animal industry on the earth, while in the meantime driving the new wave towards the changes of suitable cultivating procedures. This audit focuses on the developing mechanical headways in monitoring of animal health care for definite, exact data on productivity, and in addition physiology and prosperity. Biosensors will add to the fourth transformation in horticulture by fusing creative innovations into savvy analytic techniques that can mitigate the possibly disastrous impacts of irresistible flare-ups in cultivated creatures. There is and a comprehension of how we can uncover extra motivating force from the information delivered by human administrations and government. Far reaching measure of heterogeneous information is made by these workplaces. Nevertheless, without fitting information examination procedures these information advanced toward getting to be useless (Archenaa et al, 2015). Huge Data Analytics using Hadoop accept a convincing part in performing critical ceaseless examination on the tremendous volume of information and prepared to foresee the emergency circumstances before it happens. It portrays about the enormous information use cases in human administrations and government.

Among the most widely recognized and ceaseless issues in the social insurance framework worldwide is the swarming of crisis rooms (ER); prompting numerous genuine inconveniences. Ruler Faisal Specialist Hospital and Research Center used health care analytics techniques to recognize regions of inadequacy and propose potential upgrades to ER execution. The venture executed arrangements and monitored two pointers (ER length of stay (LOS)), reflecting proficiency, and level of patients leaving without treatment, reflecting viability of the ER. A review investigation of 26,948 ER experiences in 2014 was done in January 2015. Analytics methods were utilized to recommend process overhaul in view of results. Two suggestions were executed; a Fast-Track for bring down acuity ER patients and an interior waiting zone, for those patients who can remain vertical and extra an ER bed. 32.8% of ER patients had bring down acuity levels and under 0.5% of them were admitted to the hospital. In the wake of executing the two arrangements, the aggregate ER LOS was lessened from 20h out of 2014 to under 12h of every 2016; 40% improvement (Khalifa et al., 2016). The rates of patients left without being seen remained around 3.5%, while the rates of patients left before total treatment was fundamentally lessened from 13.5% out of 2014 to 5.5% of every 2016. Subsequently, the aggregate level of patients left without treatment was diminished from 17% of every 2014 to 9% of every 2016, with half change. Every single other factor was the same, including quantities of ER visits, Patient Acuity Level, working staff, working hours, and the tally of ER beds. Health care analytics strategies can be utilized to recognize regions of insufficiency, potential changes, and prescribe viable answers for positively improve ER execution. More arrangements ought to be inspected, for example, group triaging, patients palmar filtering, and setting a doctor in triage. Additionally, more pointers ought to be monitored, for example, the adequacy of ER treatment including the rates of revisits.

As health care in modern communication technology has been made secured using advanced security technology, a still advanced cloud computing platform (Rath et al., 2017) is currently being used for easy access of health information and better control with new encryption method. The encoded information will be secured in the cloud using provable information possession plot. To improve the assurance and security, the primal based security technique is used as a piece of securing the social protection information. The proposed framework gives better security. Using this proposed procedure, the exactness of therapeutic administrations information is ensured and conspicuous confirmation of archive piece nearness in the cloud is performed. This procedure furthermore reinforces incorporation, alteration and cancelation of the information and tries to diminish the calculation time of the server. The surprise attacks are made with encoded information and the results shows that the proposed plan beats with improved security in information accumulating when appeared differently in relation to the present methods.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Health care Monitoring frameworks are logically advancing with the possibility to change the way social insurance is as of now conveyed. A Smart Health Monitoring System, for example, the one composed in this investigation, means to computerize understanding monitoring errands. In the present investigation we have composed and built up a GSM (Global System for Mobile) based brilliant wearable framework with 3-hub Accelerometer, 3-lead ECG recording framework and constant NIBP examination framework. The gadget is fit for distinguishing sudden fall situations, heart abnormalities and hyper/hypotension; in this way rendering it suitable for constant monitoring, self-conclusion and remote finding purposes. The framework is likewise fit for sending notices to social insurance experts quickly after the discovery of any of the previously mentioned health care abnormalities with the goal that the vital advances are taken. The total examination and assessment of the signs procured from the patient have been performed utilizing a product stage: NI LabVIEW. It is a savvy, keen and wearable health care monitoring framework that is of monstrous significance not exclusively to the elderly and post-agent (Rath et al., 2017) patients, yet in addition to those dwelling in remote areas and are unequipped for bearing costly social insurance in urban hospitals or facilities.

5. CONCLUSION

In modern age, there are many prominent change in way of life of persons in metropolitan city and due to change in life style, furthermore there is possibility of more medical issues in urban zones. With the advancement of technology and its development in health care section, the lifestyle of people in metro urban regions with far reaching volume of masses is also affected by various application and organization systems. Along these lines by and by most of the urban groups are changing into savvy urban zones by getting motorized systems in each possible portion. With a target of building up another social insurance framework for ambulances in an urban city, the chapter presents a cutting edge description on health and human services control framework with enough security feature for key vehicles so the goal of safe driving uniquely in emergency situation can be satisfied in genuine sense. security solution of health care as it uses advanced technology of automating vehicles, big data analytics, swarm intelligence as well as VANET.

REFERENCES

Adame, T., Bel, A., Carreras, A., Melià-Seguí, J., Oliver, M., & Pous, R. (2018). CUIDATS: An RFID–WSN hybrid monitoring system for smart health care environments. *Future Generation Computer Systems*, *78*(Part 2), 602–615.

Archenaa, J., & Mary Anita, E. A. (2015). A Survey of Big Data Analytics in Healthcare and Government. *Procedia Computer Science*, *50*, 408–413.

Carney, T. J., & Kong, A. Y. (2017). Leveraging health informatics to foster a smart systems response to health disparities and health equity challenges. *Journal of Biomedical Informatics*, *68*, 184–189.

Farahani, B., Firouzi, F., Chang, V., Badaroglu, M., Constant, N., & Mankodiya, K. (2018). Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare. *Future Generation Computer Systems*, *78*, 659–676.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Finlay, D. (2016). Chapter Seven - Connected Health Approaches to Wound Monitoring. In J. Davis, A. McLister, J. Cundell, & D. Finlay (Eds.), *Smart Bandage Technologies* (pp. 229–244). Academic Press. doi:10.1016/B978-0-12-803762-1.00007-2

Harbouche, A., Djedi, N., Erradi, M., Ben-Othman, J., & Kobbane, A. (2017). Model driven flexible design of a wireless body sensor network for health monitoring. *Computer Networks*, *129*(Part 2), 548–571. doi:10.1016/j.comnet.2017.06.014

Hossain, M., Islam, S. R., Ali, F., Kwak, K. S., & Hasan, R. (2018). An Internet of Things-based health prescription assistant and its security system design. *Future Generation Computer Systems*, *82*, 422–439. doi:10.1016/j.future.2017.11.020

Hussain, A., Wenbi, R., da Silva, A. L., Nadher, M., & Mudhish, M. (2015). Health and emergency-care platform for the elderly and disabled people in the Smart City. *Journal of Systems and Software*, *110*, 253–263.

Khalifa, M., & Zabani, I. (2016). Utilizing health analytics in improving the performance of healthcare services: A case study on a tertiary care hospital. *Journal of Infection and Public Health*, *9*(6), 757–765.

Knoppers, B. M., & Thorogood, A. M. (2017). Ethics and Big Data in health. *Current Opinion in Systems Biology*, *4*, 53–57. doi:10.1016/j.coisb.2017.07.001

Kyriazakos, S., Valentini, V., Cesario, A., & Zachariae, R. (2017). FORECAST-a cloud-based personalized intelligent virtual coaching platform for the well-being of cancer patients. *Clinical and translational radiation oncology*.

Lepeule, J., Laden, F., Dockery, D., & Schwartz, J. (2012). Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard six cities study from 1974 to 2009. *Environmental Health Perspectives*, *120*(7), 965–970. doi:10.1289/ehp.1104660 PMID:22456598

Li, Y., Chen, C. S., Song, Y. Q., & Wang, Z. (2007). Real-time QoS support in wireless sensor networks: a survey. *IFAC Proceedings Volumes, 40*(22), 373-380.

Lloret, J., Parra, L., Taha, M., & Tomás, J. (2017). An architecture and protocol for smart continuous eHealth monitoring using 5G. *Computer Networks*, *129*(Part 2), 340–351. doi:10.1016/j.comnet.2017.05.018

Lomotey, R. K., Pry, J., & Sriramoju, S. (2017). Wearable IoT data stream traceability in a distributed health information system. *Pervasive and Mobile Computing*, *40*, 692–707.

Neethirajan, S., Tuteja, S. K., Huang, S.-T., & Kelton, D. (2017). Recent advancement in biosensors technology for animal and livestock health management. *Biosensors & Bioelectronics*, *98*, 398–407.

Olayinka, O., Kekeh, M., Sheth-Chandra, M., & Akpinar-Elci, M. (2017). *Big Data Knowledge in Global Health Education*.

Patel, P., Intizar Ali, M., & Sheth, A. (2017). On Using the Intelligent Edge for IoT Analytics. *IEEE Intelligent Systems*, *32*(5), 64–69. doi:10.1109/MIS.2017.3711653

Pramanik, M. I., Lau, R. Y., Demirkan, H., & Azad, M. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems with Applications*, *87*, 370–383. doi:10.1016/j.eswa.2017.06.027

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Qiu, H. J. F., Ho, I. W. H., Tse, C. K., & Xie, Y. (2015). A Methodology for Studying 802.11p VANET Broadcasting Performance With Practical Vehicle Distribution. *IEEE Transactions on Vehicular Technology*, *64*(10), 4756–4769. doi:10.1109/TVT.2014.2367037

Rath, M. (2017). Resource provision and QoS support with added security for client side applications in cloud computing. *International Journal of Information Technology*, *9*(3), 1–8.

Rath, M., & Panda, M. R. (2017). MAQ system development in mobile ad-hoc networks using mobile agents. In *IEEE 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, Noida, Japan (pp. 794-798).

Rath, M., Pati, B., Pattanayak, B. K., Panigrahi, C. R., & Sarkar, J. L. (2017). Load balanced routing scheme for MANETs with power and delay optimisation. *International Journal of Communication Networks and Distributed Systems*, *19*(4), 394–405.

Rath, M., & Pattanayak, B. (2017). MAQ: A Mobile Agent Based QoS Platform for MANETs. *International Journal of Business Data Communications and Networking*, *13*(1), 1–8. doi:10.4018/IJBDCN.2017010101

Rath, M., & Pattanayak, B. K. (2018). Monitoring of QoS in MANET Based Real Time Applications. In S. Satapathy & A. Joshi (Eds.), Information and Communication Technology for Intelligent Systems (ICTIS 2017) (Vol. 2, pp. 579-586). Springer. doi:10.1007/978-3-319-63645-0_64

Rath, M., Pattanayak, B. K., & Pati, B. (2017). *Energetic Routing Protocol Design for Real-time Transmission in Mobile Ad hoc Network. In Computing and Network Sustainability, Lecture Notes in Networks and Systems* (Vol. 12). Singapore: Springer.

Shah, T., Yavari, A., Mitra, K., Saguna, S., Jayaraman, P. P., Rabhi, F., & Ranjan, R. (2016). Remote health care cyber-physical system: quality of service (QoS) challenges and opportunities. IET Cyber-Physical Systems: Theory & Applications, 1(1), 40-48. doi:10.1049/iet-cps.2016.0023

Scarfò, A. (2018). The Cyber Security Challenges in the IoT Era. In M. Ficco & F. Palmieri (Eds.), Security and Resilience in Intelligent Data-Centric Systems and Communication Networks (Ch. 3, pp. 53-76). Academic Press.

Sun, E., Zhang, X., & Li, Z. (2012). The Internet of Things (IOT) and cloud computing (CC) based tailings dam monitoring and pre-alarm system in mines. *Safety Science*, *50*(4), 811–815. doi:10.1016/j.ssci.2011.08.028

Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, *55*(1), 64–79. doi:10.1016/j.im.2017.04.001

Zuhra, F. T., Abu Bakar, K., Ahmed, A., & Tunio, M. A. (2017). Routing protocols in wireless body sensor networks: A comprehensive survey. *Journal of Network and Computer Applications*, *99*, 73–97.

This research was previously published in the International Journal of Strategic Information Technology and Applications (IJSITA), 9(2); edited by Caroline Howard, Kathleen M. Hargiss, and Mohamed Amine Boudia; pages 18-34; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 5

Organizational and Social Implications

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 71
Shared Electronic Health Records as Innovation: An Australian Case

Troy Pullen
Queensland State Archives, Australia

Latif Al-Hakim
University of South Australia, Australia

ABSTRACT

Based on the Diffusion of Innovation theory, this chapter addresses the following research question: Whether factors of diffusion of innovation theory influence care providers' willingness to adopt Shared Electronic Health Records (SEHRs). Data was collected through a self-administered questionnaire distributed to over 5000 active members of the Australian Medical Association Queensland. A total of 588 valid responses were received from currently active care providers in Queensland. Multiple regression analysis and Chi-Square analysis were conducted to test the research hypotheses and answer the research question. The data revealed that while 72% of those surveyed were willing to adopt SEHRs, significant differences existed between public and private sector care providers and across the various tiers of the health system. In relation to the factors influencing future willingness to adopt, the variables comprising relative advantage were shown to have a significant impact upon future willingness to adopt shared electronic health records. The findings from this chapter will benefit those responsible for the future introduction of SEHRs, specifically by allowing policy makers to target the factors that influence care providers' willingness to adopt.

INTRODUCTION

An electronic health record provides each individual with a secure and private lifetime record of their key health history and care within the health system. The record is available electronically to authorised health care providers and in selected instances to the individual anywhere, anytime, in support of

DOI: 10.4018/978-1-7998-1204-3.ch071

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

high quality care (Health Canada, 2004, p. 3). 'The introduction of an electronic record should enable healthcare workers much greater access to timely, reliable and accurate data' (Conrick, 2006, p. 12). In addition, Conrick states that 'great gains are envisioned at the hospital interface, where quick access to a patient's medication record could be life saving'.

The development of electronic health records and the sharing of these records provides the foundation for the 'innovation' to be developed and implemented in this study. The shared electronic health record (SEHR) is not a recent innovation. The concept of SEHRs began at least 40 years ago, but the first implementations did not begin until the 1980s, and with the exception of a few countries in Europe, is still very low (Royal Australian College of General Practitioners, 2004; Schloeffel, 2004a). 'For over thirty years, there have been predictions that the widespread clinical use of computers was imminent. Yet the 'wave' has never broken' (Berner, Detmer, & Simborg, 2005, p. 3). However, due to factors such as technology interoperability problems and the previous lack of strong incentives from Governments, the widespread clinical use of computers has previously never progressed past the concept stage. The recent diffusion of SEHRs by Government's (e.g. Australia's proposed Personally Controlled Electronic Health Record (PCEHR)) has brought the innovation back into the spotlight.

Recently, much work has been undertaken to develop SEHRs to meet specific needs and requirements of various stakeholders. Countries over the past decade which invested significant resources to develop shared electronic health systems, and Standards for these systems include the UK, Netherlands, USA, Indonesia, Australia and Canada (Detmer & Gillings, 2000; Schloeffel, 2004a).

Defining Electronic Health Records

The International Standard ISO 20514 defines an electronic health record as (Schloeffel, 2004b):

A repository of information regarding the health of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorised users. It has a commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care and it contains information which is retrospective, concurrent and prospective.

Briggs (2000, p. 18) provides the following definition of electronic health records that was developed for the Australian system, HealthConnect:

An electronic longitudinal collection of personal health information, usually based on the individual, entered or accepted by health care providers, which can be distributed over a number of sites or aggregated at a particular source. The information is organized primarily to support continuing, efficient and quality health care. The record is under the control of the consumer and is stored and transmitted securely.

It is any information relating to the past, present or future physical/mental health, or condition of an individual which resides in electronic system(s) used to capture, transmit, receive, store, retrieve, link, and manipulate multimedia data for the primary purpose of providing health care and health related services.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Based on these definitions, there appears a common link. Both discuss electronic information (past and present) about individuals which can be shared between users, with the aim of improving health care and health services.

Additionally, the International Standard ISO 20514 (Schloeffel, 2004b) defines an electronic health system as:

A system for recording, retrieving, and manipulating information in electronic health records, or more broadly as the set of components that form the mechanism by which electronic health records are created, used, stored, and retrieved. It includes people, data, rules and procedures, processing and storage devices, and communication and support facilities

Early work in medical informatics focused on clinical computing with a clear goal – to improve clinical decisions and reduce medical errors – essentially though electronic access to procedure results, faster access to relevant medical information in the literature, and, from the beginning, decision support functions such as reminders and alerts (Berner et al., 2005). In the 1970s, its use centred on collating patient data and the possibility of electronically linking primary and secondary care emerged. In the '80s computers were introduced to the consulting room; and in the '90s the internet provided the potential to obtain and review useful information during the consultation (Mitchell & Sullivan, 2001; Richards, Bolton, Veale, & Quinlan, 1999). Today, in addition to providing care providers' with a number of clinical benefits such as up to date information, SEHRs are proposed to enable the secure sharing of health information between an individual's healthcare providers, whilst enabling the individual to control who can access their electronic health records (Commonwealth_of_Australia, 2011).

THE ROLE OF TECHNOLOGY

The adoption and implementation of SEHR systems consists of a range of initiatives which inevitably initially involve the introduction of new health information technology (IT) to support SEHR systems. Health IT is regarded as an essential tool for hospitals navigating the road to patient safety and implementation of SEHRs (Furukawa, Raghu, Spaulding, & Vinze, 2008). Successful implementations of health IT systems has the potential to positively influence care providers' willingness to adopt SEHRs. Research suggests that user attitudes towards IT systems are a critical factor in influencing care providers' willingness to adopt SEHRs (Castillo, Martínez-García, & Pulido, 2010)

Health IT systems are becoming increasingly critical in the daily roles of care providers'. For example, general practice is central to the coordination of health care provision with 85% of Australians interacting with a general practitioner each year (Britt, Miller & Charles 2007). It is reported that 98% of Australian general practices are computerised, with 64% recording clinical information in the form of progress notes (McInnes, Saltman & Kidd 2006). Further, a Royal Australian College of Physicians survey of 1,266 Specialists found that 97.5% of respondents had access to computers at work (Osborn et al. 2009). Today, the use of technology in health care has extended to include mobile devices such as tablets and smart phones. Mobile devices offer flexibility and ease of access to patient information (Solomon 2011). In addition to care provider use of mobile technologies, customers are using a range of mobile devices to view patient medical records such as history, medication and problem lists, and allergies (GE 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Further to the role played by health IT systems in general, the development of SEHRs has been credited with improving efficiency in health care; has the potential to increase the quality of care in hospitals, and reduce costs of clinical care (Burke & Menachemi, 2004; Darr, Harrison, Chakked, & Shalom, 2003; Lorence & Jameson, 2002).

Generally, following the implementation of SEHRs, healthcare providers are deemed to benefit via availability of information at the point of care, better information storage and retrieval, efficiency of storage space and time, more accurate records, improved drug management, improved legibility and presentation, better security and integration of clinical and administrative functions (Jefferson, 1998). Shared electronic health records enable patients to have easy-to-access information about their medical history, including medications, test results and allergies, allowing them to make informed choices about their healthcare and be active participants in their healthcare (Commonwealth of Australia 2011). Additional benefits to patients include improved prescribing of medications, improved referrals (based on the patient summary) and access to computer based patient education materials. In addition, Jefferson (1998) states that the government and healthcare system in general is believed to benefit from cost savings. There are heavy expenses nationwide with the duplication of diagnostic tests (Dearne, 2005). 'That can certainly come down if we have readily available health records so that blood tests, X-rays and MRI's do not need to be repeated by every lot of treating doctors', stated then Federal Health Minister, Mr. Abbott (Dearne, 2005).

Despite the potential benefits, access to patient records has previously often been incomplete and disorganised. This was due to the decentralized and defragmented nature of health systems (Commonwealth of Australia 2011; Kohn 2000). This is supported by Briggs (2000b), who states that, 'the majority of health care records in Australia exist as discrete paper based entities held at a variety of different locations, resulting in a fragmented picture of individuals' health needs and health histories'. Medical records can typically be stored in separate systems across GPs, hospitals, emergency departments, outpatients' clinics, diagnostic units and specialists' rooms. A clinician therefore could access only those records stored in their own files, while access for patients was extremely limited (Aap, 2004). In a study conducted by the RAND Corp. in the United States relating to the creation of an electronic health system through a unique patient identification number for every person, a number of benefits were realised. These benefits included a reduction in medical errors, increasing the overall efficiency of patient treatment, and protection of patient privacy (Carayon 2009).

The development of an integrated SEHR network has the potential to benefit patient care substantially. The Commonwealth Department of Health and Ageing (Commonwealth of Australia 2011) states that the purpose of SEHRs is to address the information fragmentation that currently exists in relation to personal health information. Further, Briggs (2000b, p.8) states that 'Electronic records and transmission of personal health information can provide a powerful tool to link the isolated islands and fragments of information that currently exist'.

A majority of American health care executives believe SEHR systems will have a positive long term financial impact on their organization, despite costs and other barriers to adoption (McGee, 2004). The key is not the technology, but how the technology can be utilised to reinvent health care (Berner et al., 2005).

The current Australian healthcare system operates in a mixed mode of using paper-based and electronic-based systems for collecting and sharing health information. A number of different eHealth applications are in wide use in a number of different areas in the health sector including patient administration systems (PAS), clinical information systems (CIS), diagnostic imaging systems, pathology systems, and practice management systems (National E-Health Transition Authority 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As clinical computer use in general practice becomes more widespread, evidence of its benefits is being accumulated. For example, 'computerised prescription software packages that are being used widely in Australian general practice and have the potential to improve the quality use of medicines through providing the general practitioner with information about medications, access to clinical guidelines and warnings about potential contraindications, adverse reactions and allergies' (Royal Australian College of General Practitioners, 2004, p. 74). Prescriptions are legible and accurate and allow audits of prescribing for individual patients and the practitioner's patient population. Individualised consumer medicine information can be produced at the time a prescription is generated (Royal Australian College of General Practitioners, 2004).

AUSTRALIAN INITIATIVES IN E-HEALTH

Early trials conducted throughout the past decade have confirmed the early evidence that health care quality can be significantly improved through health information (Detmer & Gillings 2000). Over the next two decades, e-health could deliver a positive impact on quality and access of health care. These early studies have found that these positive impacts are likely to improve cost effectiveness of health care, which should lead to better evaluation of health status and outcomes (Detmer & Gillings, 2000). In addition, Detmer and Gillings state that a benefit of allowing patient access to electronic health systems is that more knowledgeable patients will likely support rising standards of medical care and research and result in better management of the knowledge base. These findings from Detmer and Gillings in 2000 remain relevant today, with the "Concept of Operations" relating to the introduction of a personally controlled electronic health record system" (National E-Health Transition Authority 2011) containing a number of the same benefits for individuals and healthcare providers alike, stating the system will enhance the way healthcare is delivered.

Across Australia, many programs have been, and are currently being undertaken to provide not only doctors and GPs with adequate electronic medical access, but also patients. Prior to 2010, a number of SEHR initiatives have been developed. The Concept of Operations document referred to above (National E-Health Transition Authority 2011) for the personally controlled e-health records (PCEHR) system summarises the history of Australian initiatives relating to eHealth. Consideration of SEHRs in Australia started with the National Electronic Health Records Taskforce (NEHRT) in 2000, which was commissioned by the Australian Government to consider the potential for a network of electronic health records. The recommendations of the NEHRT led to the creation of the HealthConnect program and work on a range of trials on SEHRs progressed initially through HealthConnect and MediConnect programs in Tasmania, Queensland, New South Wales, South Australia, Western Australia and the Northern Territory.

In addition to the above trials, in 2004 New South Wales developed a $19.4 million pilot program called Health e-link. As part of Health e-link, clinicians and patients in NSW could retrieve information including prescriptions issued, blood test results, x-rays and hospital discharge notes for example. Tests would not have to be repeated unless absolutely needed, nor would patients and carers have to recall from memory all aspects of care they received in the past (Aap, 2004). This initiative allowed patients to access their full medical records online. In addition, health professionals – including GPs, specialists and emergency department clinicians – were able to see a patient's detailed medical history (Aap, 2004).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In 2005 it was recognised that for eHealth to progress further in Australia, key infrastructure and standards were required, and the National E-Health Transition Authority (NEHTA) was established. The development of the Personally Controlled Electronic Health Record (PCEHR) system is a result of the NEHTA, and follows on from the initiative known as HealthConnect. Below is a brief description of HealthConnect and PCEHR.

HealthConnect

HealthConnect was Australia's first health information network. The network involved the collection, storage and exchange of consumer health information via a secure network and within strict privacy safeguards to provide better integration of care and improved outcomes across the health care system (Health Connect, 2004a). The HealthConnect initiative was a partnership between the Australian, State and Territory Governments. The initiative aimed to improve safety and quality in health care by facilitating the development and establishment of a range of standards based products and services that would enable the secure exchange of information between health care providers and consumers. Privacy, security and timeliness of information flows to improve the delivery of health services were the key drivers of this initiative (Australian General Practice Network 2007).

Personally Controlled Electronic Health Records

The most recent update on Australian initiatives in SEHRs is the move towards personally controlled electronic health records (PCEHR). In April 2011, the Commonwealth Department of Health and Ageing released a 'concept of operations' for the Federal Government's $466.7 million PCEHR system set to be made available to the general public from July 2012 (ZDNet 2011).

The PCEHR system is Australia's national approach to SEHRs. The national solution extends the concepts of SEHRs and includes the required infrastructure and legislative requirements necessary to implement such a comprehensive solution. The Concept of 'Operations maps out the SEHRs function, possible information as well as proposed security and privacy settings and how the system will connect with clinical systems.

The PCEHR system focuses not only on health care administration, but also enables the secure sharing of health information between an individual's healthcare providers, while enabling the individual to control who can access their PCEHR (Commonwealth of Australia 2011). When patients visit their local GP, specialist or emergency department, the PCEHR system will allow the care provider to access patient health records at a click of a button.

The implementation of a national e-Health system addresses one of the greatest challenges faced by the Australian health system — the fragmentation of information spread across a vast number of different locations and systems (National E-Health Transition Authority 2011). The Commonwealth Department of Health and Ageing and the National E-Health Transition Authority (NEHTA) is currently working with each of the Australian State and Territory Health Departments to implement a range of foundations, including Healthcare Identifiers, Discharge Summaries and Secure Messaging, all of which will be required for the PCEHR System (National E-Health Transition Authority 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

COUNTER ARGUMENTS AGAINST THE IMPLEMENTATION OF SEHR

Despite the evidence of the benefits associated with quality information, at the point of patient care, the recognition that SEHRs could improve health care quality, reduce medical errors, and reduce health care costs is still not sufficient motivation to overcome resistance to the system's adoption. Without strong physician demand, Berner, Detmer and Simborg (2005) state that hospital and practice administrators do not see sufficient potential financial return to try to overcome this resistance either in the inpatient or outpatient setting. Claburn (2004) acknowledges that while some physicians initially resisted the technology, that is now changing. 'There's increasing emphasis on the quality of care, and increasing litigation' (Claburn, 2004). Human errors which could possibly be prevented by electronic health systems could wind up costing a hospital a million dollars in a lawsuit (Claburn, 2004). Therefore, as previously stated, a strong objective of initiatives such as HealthConnect and the current PCEHR is to encourage early uptake and installation of the system to help realise its benefits sooner rather than later. This study aims to identify those variables with the greatest potential to influence care providers' willingness to adopt SEHRs. In addition, there are three main challenges facing the implementation of SEHR systems. These are infrastructure challenges, care providers privacy concerns and consumer privacy concerns.

Infrastructure Challenges

Shared Electronic Health Record technology has a great potential to save lives as well as money, with the benefits far outweighing the initial implementation costs. However, these potential benefits can only be realised if the various systems in place across the country 'exist', and can 'talk' to each other. While rapidly improving in recent years, one primary factor identified is the lack of technology infrastructure at doctors' offices, especially small practices (Hare, 2001; McGee, 2004). Implementing this infrastructure comes at a significant cost, particularly to small practices. Secondly, the lack of interoperability between SEHRs has been a major barrier towards the deployment of shared electronic health systems (Schloeffel, 2004a). The challenge, therefore, lies in developing a national framework for infrastructure, privacy and security and common standards to support the reliable flow of information across the health sector (Health Connect, 2004b). Conrick (2006, p. 12) states that 'The introduction of standards for technology and language will enable greater connectivity and understanding between clinicians while supporting data analysis and audit'.

In order to overcome the challenges of system interoperability, as part of the development of the PCEHR program, e-health lead implementation sites have been established around Australia to deploy components of electronic health record systems, providing valuable learning to inform the implementation of the national PCEHR system (Commonwealth of Australia 2011).

Care Providers Privacy Concerns

Systems of SEHRs have traditionally proved very difficult to design and implement successfully. 'Issues to be overcome include: reassuring care providers and consumers that their privacy will be protected; the need to stick to agreed terminology; the challenges surrounding entering data on the record; providing decision support tools which health care providers value; and ensuring access only to authorized users' (Briggs, 2000, p. 21). Key questions are how much security is enough, who owns the record(s), and who

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

gets to use the data (Detmer & Gillings, 2000). There are concerns that a system such as the PCEHR system may be used to check doctors' prescribing habits or consumer entitlements like medications at concession rates (Darr et al., 2003; Merkes, 2000).

In addition to concerns that SEHRs may be used to check care providers' prescribing habits, the discoverability of data under legal proceedings encourages silence about errors committed or observed. Errors and safety issues can go undetected and unreported, both externally and within health care organisations (Kohn, 2000). Kohn (2000, p. 127) states that the 'Fear of legal discoverability or involvement in the legal process is believed to contribute to underreporting of errors'.

Further to the concerns raised above, it has been identified that there are also enormous dangers in information overload, especially in diagnosis and treatment contexts (Burgess, Gray, & Fiddian, 2004; L. Burke & Weill, 2005; Clarke, 2001; Farman, Honeyman, & Kinirons, 2003; Kerr & Norris, 2004; Lott, 1997; Murray & Lynn, 1996). The discussion relating to the amount of information available on the internet alone provides some context to the depth of information – both useful and otherwise, which must be taken into account.

Consumer Privacy Concerns

International surveys indicate that the public is particularly anxious about privacy in the context of electronic health related issue. A report by the California HealthCare Foundation found that 67% of the national respondents felt "somewhat" or "very concerned" about the privacy of their personal medical records (Bishop, Holmes, & Kelley 2005).

As relevant to the PCEHR system as to the context in which the statement was made over a decade ago, Neame and Kluge (1999) point out that 'accepting the benefit of technology is easy; offsetting it against risk is harder'. The potential for breaches of confidentiality may render patients less ready to share confidences with doctors and render doctors less likely to record them - both of which are detrimental to good clinical care (Keeley, 2000). For example, over 66% of respondents to a survey on electronic health records felt that electronic health records could reduce medical errors. Further, nearly 13% of respondents withhold personal information—such as existing health problems—which could *reduce* errors (Angst & Agarwal 2009).

DIFFUSION OF INNOVATION THEORY

This research examining the implementation of shared electronic health records (SEHRs) utilises Diffusion of Innovation theory (Rogers 1995) as a basis for its theoretical framework. Kale and Arditi (2010) state that diffusion of innovation has received significant interest from a wide range of disciplines such as management, economics, marketing, and sociology for more than four decades. This interest has generated a rich literature. Kale and Arditi (2010) conclude that the emergent picture from this literature is that the cumulative adoption of an innovation generally follows a S-shaped curve composed of: 1) an initiation and implementation phase with slow growth; 2) an adoption phase with fast growth; and 3) a saturation phase with decelerating growth (Mahajan et al. 1990).

Innovation diffusion models adopt a conceptual framework for understanding the diffusion of the innovation process (Kale & Arditi 2010). Diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system (Rogers 2003). Diffusion is

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

a communication process in which adopters persuade those who have not yet adopted to adopt (Valente 1995). Subsequently, different research approaches have been set forth in the literature to explain the observed S-shaped curve e.g., Bass 1969, Moore 1995, and Rogers 1995.

The term innovation refers to ideas, practices or objects that are perceived as new by an individual or other unit of adoptions (Rogers 1995). An innovation diffusion model focuses on describing and explaining the adoption process as a process of innovation diffusion at the aggregate level (Kale & Arditi 2010). The main objective of innovation diffusion models such as those used by Bass (1969) and Rogers (1995) is to explain or predict rates or patterns of innovation adoption over time (Mahajan et al. 1990).

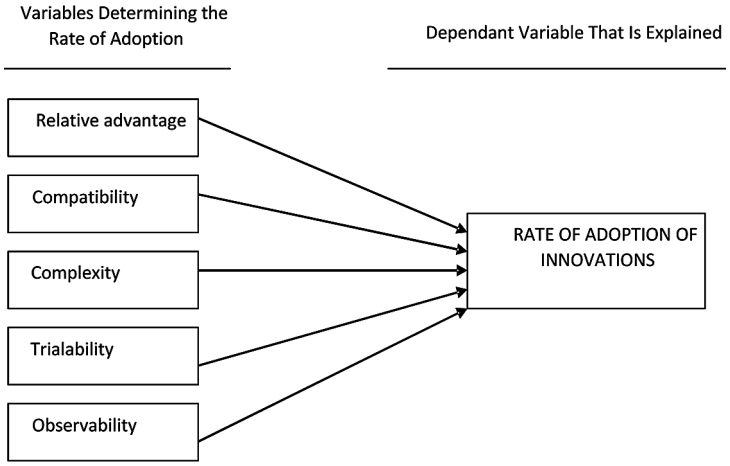
Rogers (1995) states that the characteristics of an innovation, as perceived by individuals, help to explain adoption rates. This is also consistent with other studies, which have found that the attitude of adopters to the various characteristics of an innovation influences adoption rates (Darr et al. 2003; Mahajan et al. 1990; Teng, Grover, & Güttler 2002). Variables which affect rate of adoption of an innovation, as determined by Rogers (1995), are:

1. Relative advantage;
2. Compatibility;
3. Complexity;
4. Trialability; and
5. Observability.

Rogers' framework of variables determining rate of adoption is illustrated in Figure 1.

Once the factors identified by Rogers were applied to the research, some adaptation to the theoretical framework was required in order to ensure relevance to the subject matter – adoption of SEHRs. In the context of this study, care providers' willingness to adopt SEHRs is measured. The section below outlines the rationale for focusing on care providers' willingness to adopt as opposed to measuring rate of adoption of SEHRs.

Figure 1. Rogers' framework of variables (adopted from Rogers, 1995)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Predicting Willingness to Adopt

The timing of this research is such that, at this point, SEHR use by care providers is not at a stage of diffusion where it is feasible to assess actual adoption behaviours. Similar to a study conducted by Angst and Agarwal (2009), for the most part, there are few cases in which SEHRs are stored in interoperable systems or made available via the internet to care providers or patients. Further to this, a study by Peterson et al. (2009) found that in America, the percentage of physicians using 'fully functional' electronic health record systems (e.g. incorporating extensive clinical features) only grew from three to four percent between 2006 and 2008.

Therefore, as yet, while care providers typically have not adopted the technology; they can form attitudes and beliefs about the concept of participating in adoption of SEHRs. Therefore use must be assessed through perceptual measures rather than actual opt-in behaviour. Consequently, this research will ascertain whether care providers would choose to opt-in to an SEHR system if they are given the choice in the near future. Therefore, the variable—*willingness to adopt* — is integrated into the innovation model as a means of estimating actual future behaviour.

As an additional factor impacting care providers willingness to adopt, due to the nature of implementation of Government initiatives such as that of HealthConnect or PCEHRs – the basis for a national framework of SEHRs - these projects by nature have a *mandatory* adoption procedure. Adoption of SEHRs is therefore unlike other innovations where individuals can *choose* whether to implement or not based on Rogers' variables identified above. For mandatory implementation, as the term suggests, adoption will result in *all relevant* people within the specified community adopting the innovation. Measuring rate of adoption would then become irrelevant because eventually, all members in the specified community will have to adopt the innovation. As outlined above, what may be highly variable however is adopters' *willingness to adopt*.

Willingness to adopt may impact strongly on the cost and length of the implementation process. Information provided in this Chapter has identified that the widespread clinical use of SEHRs amongst care providers is currently in its infancy. In order to promote acceptance across all care providers, considerable effort needs to be made to identify those factors which may affect willingness to adopt. By identifying those factors that have a positive or negative impact upon willingness to adopt an innovation, the effectiveness of the implementation program may be improved. This study therefore focuses upon the factors impacting upon willingness to adopt an innovation.

Research to ascertain willingness to adopt an innovation will invariably require some adaptation to the variables Rogers presented for measuring rate of adoption. The first variable Rogers identified as affecting an innovation's rate of adoption is relative advantage. Relative advantage gauges the degree of advantage gained from implementation of an innovation, compared to the previous system/s used. If an innovation has a perceived advantage or disadvantage over what it is designed to replace, those adopting the innovation will be either more, or less willing to adopt based on the degree of advantage the new innovation has. For example, if an innovation produces a cost saving to an organisation of 10 percent, those adopting would be more *willing*. Therefore, relative advantage can still be considered an important variable which may impact upon willingness to adopt SEHRs.

Compatibility is the second variable used by Rogers to analyse rate of adoption. As outlined by Rogers, compatibility can be viewed in terms of being consistent with existing values and procedures. Compatibility in relation to the current study can be viewed from a number of perspectives. Firstly, compatibility may be viewed from the perspective of technology usage. Those who indicated that they

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

have a high level of technology usage may be likely to have positive attitudes towards technology, and be more likely to find SEHRs more compatible with current systems. Consequently, based on positive compatibility, care providers may be more willing to adopt. Additionally, compatibility can potentially challenge existing values. If innovations challenge deeply embedded work patterns or subcultures, care providers may be reluctant to make the innovation part of their work routines (Buell 2009; Darr et al. 2003; Reid 2010; Wu et al. 2010). Effective communication can enable or disable the adaptiveness needed to navigate change successfully (More & McGrath, 2002). Therefore, compatibility with existing values or procedures could be considered relevant in influencing a care provider's willingness to adopt SEHRs.

The third variable identified by Rogers as affecting rate of adoption is complexity. Again, like relative advantage and compatibility, complexity is a factor which could impact upon an individual's willingness to adopt. For example, if a system of SEHRs is perceived to pose complexities in terms of: (a) meeting regulatory and legal requirements; or (b) technical difficulties, care providers may be less willing to adopt the new system (Wu et al. 2010).

The fourth variable which Rogers presents as impacting upon rate of adoption is trialability. The notion of trialability refers to the ability to trial an innovation before making the eventual decision to implement or not to implement an innovation. However, in the case of mandatory implementation, individuals have no individual ability to trial an innovation and then choose whether to adopt or not. Therefore, trialability could be considered as not directly influencing an individual's willingness to adopt. However SEHRs, as explored earlier in this chapter, have been trialled in the national SEHR initiative of HealthConnect and other smaller pockets within individual states of Australia. These trials have taken place over the past several years, with the aim of ensuring the system runs reliably and can be efficiently implemented nation wide. Therefore, those not involved in the HealthConnect or PCEHR trials have the capacity to learn about the results of the trials, and analyse feedback from those participating in trials. Therefore, while individuals themselves cannot trial SEHRs, they do have the capacity to analyse results of trials conducted by the Government. The results of these official trials may therefore either positively or negatively, influence care providers' willingness to adopt SEHRs.

Observability is the final variable which Rogers uses to measure rate of adoption of an innovation. Observability relates to those involved in the adoption process being able to observe the results of an innovation. Observability in this study is similar in nature to the previous variable, trialability. While an individual cannot observe the results of a trial of SEHRs in their own organisation before deciding to adopt, they have the capacity to observe the results gained from official trials. In the context of this study, observability may influence care providers' willingness to adopt from analysing the willingness to adopt across various tiers of medical professionals or across public and private hospitals.

Determining the factors which have the greatest potential for influencing willingness to adopt innovations would be considered highly desirable. In the case of this study, the innovation is SEHRs. The desirability of being able to identify factors likely to impact upon willingness to adopt could have significant potential for policy makers, both within and beyond the concept of SEHRs. By targeting those factors which may influence willingness to adopt, the adoption process can become a more efficient, streamlined process which meets the demands of the adopters'. The current study is specifically targeting care providers' perceptions of relative advantage compatibility, complexity and observability via analysing the willingness to adopt SEHRs. Supporting the focus upon relative advantage, compatibility and complexity, Carter and Bellanger (2005) state that after an extensive literature review, Tornatzky and Klein (1982) conclude that relative advantage, compatibility and complexity are the most relevant

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

characteristics to adoption research. Teng et al. (2002) supports this position by stating that the most important characteristics which influence adoption of technological innovations (e.g. SEHRs) are: relative advantage, compatibility and complexity. However, we are considering observability from different aspects, that is, through analysing the responses for the willingness to adopt across various tiers of medical professionals as well as across public and private hospitals.

RESEARCH METHODOLOGY

Target Population

The population used for this research was all currently active Health care providers registered with the Australian Medical Association Queensland (AMAQ). This population consisted of approximately 5013 members, located throughout Queensland. The roles of these care providers' could be broken down into the following categories: Specialists; General Practitioners; Residents; Registrars; Salaried Specialists and Senior Medical Officers. These members were assumed to have some degree of awareness of SEHRs. AMAQ member knowledge of SEHRs may have possibly been gained from potential involvement in trials of the HealthConnect system, interaction with colleagues involved with HealthConnect, and/or as a result of extensive levels of regulation developed for each of the respective states (Briggs 2000; Health Connect 2004a; Health Connect 2004b; Parker, Hewitt, & McWilliam, 2004).

Data was collected through a mail survey sent to all active members of the AMAQ. The survey was distributed by the AMAQ mail house, but returned directly to USQ to protect confidentiality of AMAQ members.

Survey Design

Ideally interviews would have been carried out, but this was not possible owing to budget, time and geographic constraints. The questionnaire was organized into the following five themes;

Part A: Demographics,
Part B: Technology,
Part C: Patient Care,
Part D: Legislation,
Part E: Shared Electronic Health Records.

These themes cover the many facets of SEHRs, and provide data about the perceived factors which influence care providers' willingness to adopt SEHRs. Questions relating to demographic detail were added in order to assist in data collection and to form the basis of comparisons across business sectors and regional differences.

A number of methods were used in the survey design to assist in building the response rate. In order to increase the response rate, Nicastro (1999) states that the covering letter should provide details relating to who should complete the questionnaire and suggests that it should be attached to the survey. It was felt to be important that the covering letter illustrated the importance of the use of technology to facilitate

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the management, and sharing of information in order to aid efficient patient care. Zikmund (2003) additionally suggests that the covering letter should explain that the questions are simple and to the point, and that a follow up email or letter to remind the respondent to complete the survey would be sent.

Data Preparation and Techniques Used

Prior to any analysis, variables were recoded for consistency and applicability. The data was entered into SPSS, and was checked for accuracy by running frequency distributions and any out of range values were checked. Whilst some cases showed the presence of outliers, none were deemed to be inconsistent and therefore all were retained in the analysis. SPSS was also used to check for missing values. The percentage of missing values for all variables was small.

Responses to interval scale questions such as those measured on a 5-point Likert scale were recoded so as the positive responses were assigned a higher number compared to negative responses. For example, 'Very poorly' was recoded as a '1', 'Poorly' recoded as '2', 'Adequate' remained as '3', 'Very well' was recoded as '4', 'Excellent' recoded as '5'.

It was considered appropriate to apply parametric testing procedures on the 5-point Likert scale in which this hypothesis is based. A review of the literature found that Likert scales are commonly treated as interval data for testing purposes, even though the data is ordinal in nature (Goldstein & Hersen, 1984; Jaccard & Wan, 1996; Shaughnessy & Zechmeister, 1990; Sisson & Stocker, 1989), facilitating parametric data analysis. The application of parametric tests to Likert scales has been validated by the fact that these tests do not affect Type I and Type II errors dramatically (Jaccard & Wan, 1996).

Hair et al (1998, p. 14) defines multiple regression as 'the appropriate method of analysis when the research problem involves a single metric dependent variable presumed to be related to two or more metric independent variables'. In addition to the definition provided by Hair et al, Zikmund (2003, p. 576) states that multiple regression analysis is 'an analysis of association in which the effects of two or more independent variables on a single, interval scaled or ratio scaled dependant variable are investigated simultaneously'.

Chi-Square analysis was used to examine the strength of the association between collected data and the specified distribution. Specifically related to this study, Chi-Square tests are used to analyse any statistically significant difference in willingness to adopt SEHRs between the different tiers of the health system within Queensland, and to test for any statistically significant differences in willingness to adopt SEHRs between care providers from public and private sector health facilities within Queensland.

Research Hypotheses

This research considers the shared electronic records as innovation and raises the following questions: Which of factors of diffusion of innovation theory influence care providers' willingness to adopt shared electronic health records (SEHRs). The study considers four factors of diffusion of innovation theory. These factors are relative advantage, compatibility, complexity and observability. The study excludes the trialability for the reasons stated earlier. In order to answer the research question, four hypotheses are established. These are as follows:

Hypothesis 1: That the variables comprising relative advantage have the potential to influence willingness to adopt shared electronic health records;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hypothesis 2: That the variables comprising compatibility with existing systems and practices have the potential to influence willingness to adopt shared electronic health records;

Hypothesis 3: That the variables comprising complexity have the potential to influence willingness to adopt shared electronic health records;

Hypothesis 4: That the willingness to adopt shared electronic health records varies across the three tiers of the health system within Queensland (observability 1);

Hypothesis 5: That the willingness to adopt shared electronic health records varies across public and private sector health facilities within Queensland (observability 2).

DATA ANALYSIS: DESCRIPTIVE ANALYSIS

Response Rate

A total of 588 valid responses (complete responses from those currently active members registered with the AMAQ) were received. The valid response rate was 11.73%.

Principal Role as the Respondents

Based on valid responses, 214 (36.4%) respondents indicated that their principal role was that of a Specialist; 202 (34.4%) operated as a General Practitioner; 101 (17.2%) described their role as that of Resident or Registrar; 43 (7.3%) were a Salaried Specialist and 16 (2.7%) respondents indicated that their primary role as a care provider was acting as a Senior Medical Officer (refer Table 1). A small number of responses (2%) were received for 'Other'. These responses included 'Medical Assessor', 'Medical Research' and 'Private Emergency Department'.

Sector Operated in by the Respondents

Table 2 provides an overview of the sectors operated in by respondents. Almost half of the respondents (49.9%) indicated that they operated within the Private sector (Table 2). One third of the respondents (33.6%) indicated that they operated solely within the Public sector, and 16.5% respondents operated within both the public and private sectors.

The greater number of respondents from the private sector could potentially impact trialling methods of SEHRs, and eventual willingness to adopt. Adoption of SEHRs has historically been trialled more widely within the public sector. The increased exposure to SEHRs within the public sector could potentially lead to a greater willingness to adopt from public sector employees.

Hospital Size

The majority of the respondents (45.4%) indicated that they operated within a large hospital. A large hospital is defined as one which has more than 200 beds. One hundred and seventeen (21.8%) of the respondents indicated that they operated within a small hospital (refer Table 3)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Principal role of the care providers

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Specialist	214	36.4	36.4	36.4
	Gp	202	34.4	34.4	70.7
	Resident and Registrar	101	17.2	17.2	87.9
	Salaried Specialist	43	7.3	7.3	95.2
	Senior Medical Officer	16	2.7	2.7	98.0
	Other	12	2.0	2.0	100.0
	Total	588	100.0	100.0	

Table 2. Sector which care providers operate within

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Private	293	49.8	49.9	49.9
	Public	197	33.5	33.6	83.5
	Public and Private	97	16.5	16.5	100.0
	Total	587	99.8	100.0	
Missing	System	1	.2		
Total		588	100.0		

Table 3. Size of hospital

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Small	117	19.9	21.8	21.8
	Large	244	41.5	45.4	67.2
	N/A	176	29.9	32.8	100.0
	Total	537	91.3	100.0	
Missing	System	51	8.7		
Total		588	100.0		

Description of Work Category

Respondents were asked to best describe the type of work organisation they were involved with. Two hundred and seventy nine respondents (47.6%) indicated that they worked within a Hospital/Medical Care environment (Table 4). The next largest respondent group was made up of those working within Group Practice (26.8%), with the self employed professional group the third most prominent category (18.6% of respondents). Additionally, 26 respondents (4.4%) indicated that they worked within two or more categories simultaneously.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Best description of work category

Category		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Hospital/Medical Centre	279	47.4	47.6	47.6
	Group Practice	157	26.7	26.8	74.4
	Ambulatory care clinic	8	1.4	1.4	75.8
	Long term care/rehabilitation centre	2	.3	.3	76.1
	College/University	2	.3	.3	76.5
	Consulting firm	3	.5	.5	77.0
	Self employed professional	109	18.5	18.6	95.6
	Other work setting	26	4.4	4.4	100.0
	Total	586	99.7	100.0	
Missing	System	2	.3		
Total		588	100.0		

Computer Applications Used

The medical community represents a diverse community. Within different roles and environments, many types of computer applications are likely to be used. Of interest to this study, current usage of technology may influence perceptions of SEHR compatibility with current systems. Respondents were asked to identify those applications which they commonly use in their daily tasks (Table 5).

The three most commonly used computer applications indicated by respondents were:

- Administrative (481 responses);
- Database searching (442 responses); and
- Colleague interaction (395 responses).

To positively influence care providers' perceptions of SEHRs and their willingness to adopt, the most commonly used applications identified by care providers' in their daily activities should be identified as strengths of SEHRs. Therefore, the implementation of SEHRs could therefore target improvements in the way in which care providers' interact with technology applications in areas such as administrative use, database searching, and colleague interaction.

Level of Technology Change over the Past Five Years

The technology section of the survey initially identified care providers' perceptions in regard to the level of change in technology throughout the past five years. Table 6 illustrates that the majority of respondents (79.2%) identified either an increase (42.4%) or a large increase (36.8%) in the level of technological change over the past five years.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Computer applications used by the respondents

Technology Type	Frequency	Percent
Administrative	481	82.1
Database searching	442	75.4
Colleague interaction	395	67.4
Prescription Use	344	58.7
Statistic collection	242	41.3
Decision Support	219	37.4

Table 6. Level of change in technology over the past five years

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Large Increase	215	36.7	36.8	36.8
	Increase	248	42.3	42.4	79.1
	Steady	116	19.8	19.8	99.0
	Decrease	1	.2	.2	99.1
	No Technology Usage	5	.9	.9	100.0
	Total	585	99.8	100.0	
Missing	System	1	.2		
Total		586	100.0		

Ease of Use of Technology

Ease of use of technology could be considered a significant factor in determining individuals' acceptance of technology. Technology which is easy or common to use, such as most typical Windows packages, are often accepted with greater ease. This theme was certainly demonstrated by responses to this question in the survey. The vast majority of respondents (96.6%) indicated that ease of use of technology would positively influence acceptance of technology (refer Table 7).

Table 7. Respondents' answers to the question: 'Do you perceive 'ease of use' technology as a factor which would influence acceptance of technology?'

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	564	96.2	96.2	96.2
	No	20	3.4	3.4	100
	Total	584	99.7	100.0	
Missing	System	2	0.3		
Total		588	100.0		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Level of Input towards Technology Use within Organisation

The next section of the technology area sought to identify whether respondents had any input into the level of technology usage within their organisations. Prior input in the development or implementation of technology could potentially influence willingness to adopt SEHRs. Approximately half of the respondents (52.7%) indicated that they had had some degree of input into the level of technology usage within their organisation (Table 8).

Perceptions of Usefulness of Technology

This component of the technology section identified personal perceptions of the use of technology for work purposes. Table 9 indicates that the majority of respondents (87.9%) rated technology, in varying degrees, as being useful in performing work duties.

Based on the results of the above table, in order to positively influence care providers' willingness to adopt technology - specifically SEHRs, the technology needs to be perceived as a tool which will enable improved work performance. These improvements must be able to be clearly defined through a number of key benefits – namely compatibility with existing work practices and applications, and relative ease of use.

Table 8. Respondents' answers to the question: 'Have you had any input into the level of technology used within your organisation?'

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	307	52.4	52.4	52.7
	No	276	47.1	47.3	100
	Total	583	99.5	100.0	
Missing	System	3	.5		
Total		588	100.0		

Table 9. Usefulness of technology for work

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Complex and Cumbersome	67	11.4	11.5	11.5
	Useful in performing daily duties	376	64.2	64.4	75.9
	Assists in selected duties	137	23.4	23.9	99.3
	Others	4	.7	.7	100.0
	Total	584	99.7	100.0	
Missing	System	2	.3		
Total		586	100.0		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ability of Technology to Provide Autonomy

Respondents were asked to identify their perceptions regarding the impact of technology, specifically whether the use of technology facilitates professional autonomy. Approximately two-thirds of respondents (63%) indicated that technology did provide some degree of professional autonomy (Table 10). These findings were verified with a follow up question which was reworded, asking respondents to consider whether technology threatens work autonomy. Consistent with the first response, approximately three quarters of respondents (78.4%) indicated that they did not perceive technology as threatening their work autonomy.

The next question identified care providers' perceptions as to whether technology could lead to managerial control. Again, consistently with the previous two questions, approximately two-thirds of respondents (66.8%) identified that they did not perceive technology as leading to managerial control over their work (Table 11).

The results indicate that care providers do not appear threatened by technology, in that the majority of respondents indicated that the use of technology does not threaten work autonomy or lead to managerial control, but rather helps facilitate professional autonomy.

Perceived Impact of Technology on Care Provider Performance

Respondents' were asked to provide an indication as to whether they perceived technology as improving their performance as a care provider. Once again, results were consistent with previous responses, and over three quarters of respondents (77.7%) indicated that they perceived the use of technology as improving their performance as a care provider (Table 12).

Table 10. Respondents' answers to the question: 'Does the use of technology provide any professional autonomy?'

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	366	62.5	62.5	62.5
	No	215	36.6	46.6	100
	Total	581	99.1	100.0	
Missing	System	7	.9		
Total		588	100.0		

Table 11. Respondents' answers to the question: 'Does the use of technology lead to managerial control over your work?'

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	195	33.1	33.2	33.2
	No	390	36.6	66.8	100
	Total	585	99.7	100.0	
Missing	System	3	.4		
Total		588	100.0		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 12. Respondents' answer to the question: 'Does the use of technology improve your performance as a care provider?'

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	454	77.5	77.7	77.7
	No	130	22.2	22.3	100
	Total	585	99.7	100.0	
Missing	System	2	.3		
Total		586	100.0		

Types of Technology Used within Daily Practice

After establishing attitudes and perceptions towards technology and the role it plays within the daily environment of the care provider, the next section of the survey identified those common activities performed by care providers' which utilised technology. The activity which care providers' identified as most commonly using technology for was data entry (77.8%). The use of technology for 'reference' and for 'treatment' use were the second and third most commonly identified uses of technology in patient care (67.7% and 60.5% respectively, refer Table 13).

As indicated above, there were a number of responses to 'Other' for this section. The most common response category for 'Other' was the use of technology to identify patient trends/statistical collection/ diagnosis investigation (11 responses).

DATA ANALYSIS: HYPOTHESES TESTING

Multiple Regression Analysis

The results of the multiple regression analysis for Relative advantage, Compatibility and Complexity are shown in Table 14.

Table 13. The most common daily uses of technology

Technology Type	Frequency	Percentage
Data entry	455	77.8
Reference	396	67.7
Treatment	354	60.5
Clinical visits	289	49.4
Decision Support	287	49.1
Prescription entry	282	48.2
Advanced use	104	17.8
Other	43	7.0

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 14. Model summary for regression analysis: Hypotheses one, two and three

Dimension	R Square	Regression Coefficient (B)		Sig
Relative advantage	.108	Perceived benefits of technology	.334	.000
		Perceived advantages of SEHRs		
Compatibility	.239	Perceived potential to improve communication between care providers		.000
		Perceived compatibility of SEHRs with care providers values relating to patient care		
Complexity	.008	Perceived complexity relating to information and records management		.037
		Perceived complexity relating to compliance		

Hypothesis 1: Relative Advantage

The value of R Square (10.8%) for Relative advantage was statistically significant (df = 2, 551; p = <.001). The regression model suggests that 10.8% of the variation in care providers' future willingness to adopt SEHRs could be explained from the variables comprising Relative advantage alone. Although the R Square value may be seemingly low (the large sample size can decrease the R Square value), it is still a statistically significant result, though some caution is needed when interpreting the results.

An examination of the beta coefficients showed that the variables comprising perceptions of the benefits of technology (Beta = .334; p = <.001) and perceptions of the advantages of SEHRs (Beta = .334; p = <.001) were significant predictors for care providers' future willingness to adopt SEHRs, indicating a positive linear relationship between Relative advantage and willingness to adopt.

An analysis of the findings also revealed that the perceived benefits of technology had a potential influence on care providers' future willingness to adopt SEHRs compared to the perceived advantages of SEHRs, with $F > 34$ and $\rho \leq 0.001$ (Table 15). While caution is needed giving the low R square ($R^2 = 0.108$), hypothesis 1 is supported.

Hypothesis 2: Compatibility

The value of R Square (23.9%) for Compatibility was statistically significant (df = 2, 551; p = <.001). The regression model suggests that 23.9% of the variation in care providers' future willingness to adopt SEHRs could be predicted from the variables comprising Compatibility alone. Similarly to the findings from the multiple regression analysis of Relative advantage, the results for Compatibility need to be interpreted with caution given the low R Square value.

An examination of the beta coefficients showed that the variables comprising perceived potential to improve communication between care providers (Beta = .501; p = <.001) and perceived compatibility of SEHRs with care providers values relating to patient care (Beta = .334; p = <.001) were significant predictors for care providers' future willingness to adopt SEHRs, indicating a positive linear relationship between Compatibility and willingness to adopt SEHRs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 15. Summary of hypothesis testing for Relative advantage

Hypothesis	R Square	Df	F	Sig Level	Accept/ Reject
H_1 That the variables comprising relative advantage have the potential to influence willingness to adopt shared electronic health records.	.108	2,551	34.539	<.001	Accept

Further analysis of the findings (Table 16) revealed that perceived compatibility of SEHRs with care providers' value had a significant influence (ρ =0.001) on care providers' willingness to adopt SEHRs. Accordingly, hypothesis 2 is supported.

Hypothesis 3: Complexity

Although the regression model suggests that less than one percent of the variation in care providers' future willingness to adopt SEHRs could be explained from the variables comprising Complexity alone, the value of R Square (<1%) for Complexity was statistically significant (df = 2, 551; p = .037).

An examination of the beta coefficients showed that the variables comprising perceived complexity relating to compliance (Beta = .113; p = <.015) were statistically significant predictors for care providers' future willingness to adopt SEHRs. Perceived complexity relating to information and records management requirements was not a significant predictor (Beta = .006; p = .903).

More detailed analysis of the findings revealed that the complexity has very low R square ($R^2 = 0.12$) with insignificant value (ρ > 0.037). Accordingly, hypothesis 3 is rejected

Hypothesis 4: Willingness to Adopt Across Three Tiers

The fourth hypothesis tests that willingness to adopt shared electronic health records varies across the three tiers of the health system within Queensland.

Table 16. Summary of hypothesis testing for Compatibility

Hypothesis	R Square	Df	F	Sig Level	Accept/ Reject
H_2 That the variables comprising compatibility with existing systems and practices have the potential to influence willingness to adopt shared electronic health records.	.239	2,551	87.886	<.001	Accept

Table 17. Summary of hypothesis testing for Complexity

Hypothesis	R Square	Df	F	Sig Level	Accept/ Reject
H_3 That the variables comprising complexity have the potential to influence willingness to adopt shared electronic health records.	.012	2,551	3.313	.037	Reject

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chi-Square analysis was chosen to test for significant difference in future willingness to adopt SEHRs between the three independent groups of the health system (i.e. General Practitioners; Specialists; and Hospital staff) so as to examine the strength of the association between collected data and the specified distribution. The p-value is the probability of observing a Chi-square statistic at least as large as the one actually observed, given that there is no association between a variable of data occurring by chance (Zikmund 2003). Results of the Chi-Square test for hypothesis four are presented below in Table 18.

The output in Table 17 shows that the value of chi-square (χ^2) = 27.275 and the degrees of freedom on which this was based, was 8. These results show that there is sufficient statistical evidence to indicate that willingness to adopt shared electronic heath records varies across the three tiers of the health system within Queensland, (χ^2= 27.275, df = 8, p=.001). Hence, there is sufficient statistical evidence to support hypothesis four.

A comparison of means across the tiers of the health sector demonstrates that Hospital staff had the highest willingness to adopt, with a mean score of 3.65 (Table 19). Specialists had the next highest willingness to adopt, with a mean score of 3.12, and GPs had a mean score of 3.09.

Hypothesis 5: Willingness to Adopt Across Health Facilities

The fifth hypothesis tests that willingness to adopt shared electronic health records varies across public and private sector health facilities within Queensland.

This hypothesis was tested using Chi-Square analysis to test for any statistically significant results relating to willingness to adopt SEHRs between public and private sector health facilities within Queensland. The Chi-Square analysis test assumes that there is no significant difference between the expected and the observed results (Garson 2009). Results of the Chi-Square test for hypothesis five are presented below in Table 20.

Table 18. Chi-Square Test on difference between different tiers of the health system and future willingness to adopt SEHRs

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	27.275[a]	8	.001
N of Valid Cases	540		
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.50.			

Table 19. Comparison of mean willingness to adopt across the tiers of the health system

Role	Mean	N	Std. Deviation
GP	3.09	191	1.184
Specialist	3.12	199	1.126
Hospital Staff	3.65	150	1.081
Total	3.26	540	1.159

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 20. Chi-Square Test on difference between public and private sectors for future willingness to adopt SEHRs

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	39.717[a]	4	.000
N of Valid Cases	465		
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.10.			

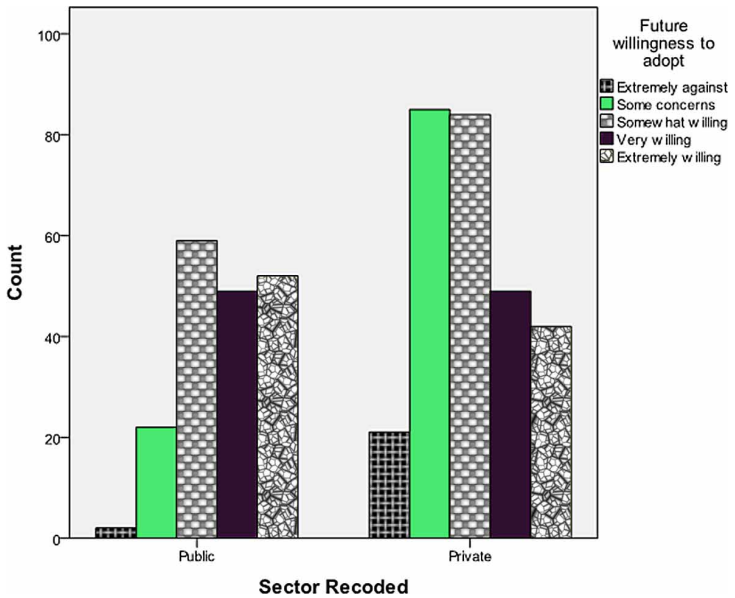
The output in Table 20 shows the value of chi-square (χ^2) = 39.717 and the degrees of freedom on which this was based, was 4. These results show that there is sufficient statistical evidence to indicate that willingness to adopt SEHRs with the sector the care provider works in within Queensland, (χ^2= 39.717, df = 4, p=.000). Hence, there is sufficient statistical evidence to support hypothesis five. Figure 2 provides a graphical representation of the public and private sectors future willingness to adopt SEHRs.

In both the public and private sectors, 'Extremely against' was recorded the least. Within the private sector, 'Some concerns' was higher than 'Somewhat willing', 'Very willing' and 'Extremely willing', but in public sector respondents, it was less. Public sector respondents were over represented in the 'Extremely willing' category, with 52 actual responses (compared to the expected total of 37.2. The greatest difference existed within the 'Extremely against' category, where 91.3% of the responses consisted of private sector respondents.

A comparison of mean scores for future willingness to adopt between the public and private sector (Table 21) finds that the public sector has a greater willingness to adopt SEHRs, with a mean score of 3.69 (compared to 3.02 for the private sector).

In conclusion, the results for testing the five hypotheses are as follow (Table 22).

Figure 2. Bar chart illustrating difference between public and private sector future willingness to adopt SEHRs



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 21. Comparison of mean willingness to adopt between public and private sector

Sector Recoded	Mean	N	Std. Deviation
Public	3.69	184	1.044
Private	3.02	281	1.174
Total	3.29	465	1.170

Table 22. Results of testing the five hypotheses

No.	Hypothesis	Accept/ Reject
1	That the variables comprising relative advantage have the potential to influence willingness to adopt shared electronic health records	Accept
2	That the variables comprising compatibility with existing systems and practices have the potential to influence willingness to adopt shared electronic health records	Accept
3	That the variables comprising complexity have the potential to influence willingness to adopt shared electronic health records	Reject
4	That the willingness to adopt shared electronic health records varies across the three tiers of the health system within Queensland (observability 1);	Accept
5	That the willingness to adopt shared electronic health records varies across public and private sector health facilities within Queensland (observability 2).	Accept

VALIDITY TESTING

A number of tests of the validity of the model for Relative advantage, Compatibility and Complexity were investigated as follows:

- The Durbin-Watson statistics (Table 23) for Relative advantage, Compatibility and Complexity indicated that there was no significant first-order serial correlation among the residuals (Zikmund 1991).
- A normal plot of regression standardised residuals for the dependent variable indicated a relatively normal distribution (Figures 3 and 4 show the normal plot of regression standardised residuals for relative advantage and compatibility).
- A scatterplot of residuals against predicted showed no clear relationship between residuals and predicted values consistent with the assumption of linearity (Coakes and Steed 1996).

The outcome of the three tests validate our results regarding hypotheses 1, 2 and 3.

Table 23. Durbin-Watson statistics for Relative advantage, Compatibility and Complexity

Variable	Durbin-Watson Statistic
Relative advantage	2.134
Compatibility	2.082
Complexity	2.106

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Regression standardized residuals for Relative advantage

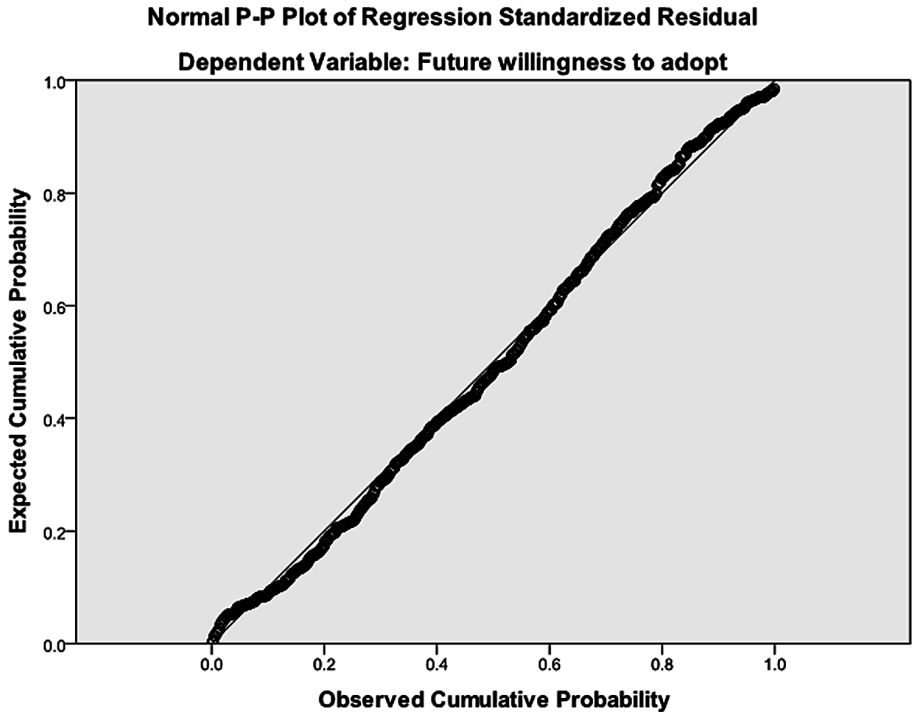
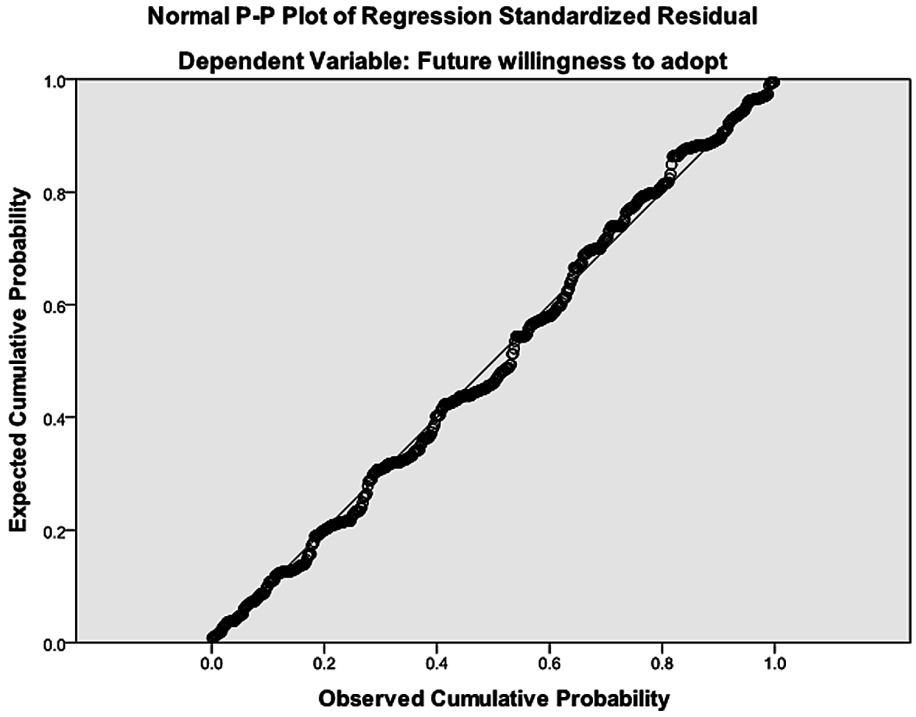


Figure 4. Regression standardized residuals for Compatibility



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

STUDY LIMITATIONS

Like other research works, this study has several limitations, including the timing of the survey, low response rate and conducting the study in Queensland only.

The questionnaire for this study was administered in 2005, this was before any widespread trialling (or implementation) of SEHRs had commenced within Australia. As a consequence, it is likely that the responding Australian Medical Association Queensland (AMAQ) registered health care providers may have had a limited understanding of SEHRs. Because of the potential for limited understanding, respondents may not have been able to provide valid perceptions of the Relative advantage, Compatibility and Complexity aspects of adopting SEHRs. Given that the findings of the study are not what the literature suggested, the study may have been ambitious given the possible restrictions (e.g. potentially as a result of the limited experience of care providers' with SEHRs at the time of survey administration) related to care providers' offering data on perceptions relating to the various aspects of SEHRs.

Considering the attempt of the Federal Government to implement SEHR technology through their Personally Controlled Electronic Health Record (PCEHR) scheme by July 2012, views of care providers' could potentially have changed due to the upcoming implementation. Indeed, future research to compare care providers' perspectives of SEHRs given the national implementation of the PCEHR system could prove valuable to compare any changes.

In addition to the development and implementation of national programs such as the PCEHR system, a cultural change is occurring within the Australian public sector. Recently, the Commonwealth government recognised that, to be more relevant to its stakeholders, it needs to foster a culture of openness. To that end, a 'Declaration of Open Government' has been developed (Department of Finance and Deregulation 2010). Community expectations about improved access to government records and information have implications for how government manages public sector information. The move towards open government recognises that the value of public sector information as a national strategic resource can be enhanced when it is openly accessible and easily reusable.

Specifically within Queensland, reforms include the change from 'Freedom of Information' to 'Right to Information', moving to a push model for data publication. These changes have resulted in the introduction of the *Right to Information Act 2009* and *Information Privacy Act 2009*, replacing the previous *Freedom of Information Act 1992.* These changes in philosophy towards information resources and an open government may result in changes to responses to this study, and perhaps reinforce the importance of the need for effective information and records management practices. Supporting and maintaining the link between good information and records management practices and ongoing open access to government information is critical, and something which SEHRs have great potential in facilitating.

The data collection instrument utilised for this study could also be considered a limitation. All information including future willingness to adopt SEHRs were self reported by respondents. While self administered questionnaires are considered a generally accepted measure of administration (Zikmund 2003), future replications of the study could use other objective measures to obtain more valid and reliable results. Such methods could include interviews and observations.

The administration method of this questionnaire could also be considered a limitation. While distribution of the survey by mail can be considered a typical way to distribute surveys (Zikmund 2003), there are a number of limitations. For example, respondents could not be identified. Because respondents were not able to be identified, this made it impossible to test for non-response bias, which could have

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

affected results. The distribution of the survey by mail also does not ensure that the intended recipient completed the survey. For example, if the recipient was different to the intended person, their views may be different and therefore results could be inaccurate.

A number of techniques were utilised in this study with the aim of reducing the potential for non-response bias and to ensure the most appropriate person completed the survey. For example, a cover letter outlining potential benefits and relevance to the care provider was developed and attached to the survey so as to influence the care provider to complete the survey. The large sample size, covering all active members of the AMAQ, was also used in an attempt to ensure responses representing the views of the care provider community were attained.

A final potential limitation of this study is the survey of only Queensland care providers' on a topic that is not limited in scope to Queensland only, but contains implications for care providers' across Australia (given the planned introduction of the PCEHR System from July 2012). Given that the role of care providers' is consistent across Australia, it is not envisaged that the focus on Queensland based care providers' would create the potential for high levels of bias.

CONCLUSION

This study sought to investigate three key factors which may impact upon care providers' future willingness to adopt shared electronic health records (SEHRs) i.e. care providers' perceptions of relative advantage, care providers' perceptions of compatibility with existing values, systems and practices, and care providers' concerns about complexity. The study employs an adaptation of Rogers' (1995) Diffusion of Innovation (DOI) theory.

Despite the benefits of SEHRs, their widespread implementation has been extremely slow to date. Previous implementation attempts have been ad hoc, and have not adequately addressed key care provider and patient concerns such as those dealing with privacy and infrastructure requirements. Being able to identify factors which may explain the slow uptake of SEHRs could be considered highly desirable, both within and beyond the concept of SEHRs. By targeting those factors which may influence willingness to adopt, the adoption process could become a more efficient process targeted to address the concerns of the adopters.

To gain a deeper understanding of the issues involved and to explore the study's research questions fully, the quantitative phase of the research involved a self administered questionnaire involving all active Australian Medical Association Queensland (AMAQ) registered health care providers from the public and private sector. The goal of the questionnaire was to identify from a care providers' perspective, factors which have the most significant impact upon future willingness to adopt SEHRs. The questionnaire was sent to currently active health care providers registered with the AMAQ. This population consisted of approximately 5013 care providers, located throughout Queensland. A total of 588 valid responses were received from currently active care providers registered with the Australian Medical Association of Queensland (AMAQ) for the study (valid response rate of 11.73%).

The majority of respondents (36.4%) indicated that their principal role was that of a Specialist. The Brisbane City region had the greatest representation of the 14 distinct demographic areas identified by the AMAQ (25.3%). Approximately half of all respondents operated within the private sector.

A number of issues related to technology, specifically respondents' general attitudes towards the use of technology within their roles as care providers were examined. The majority of care providers' (79.2%) indicated a large to very large change in the level of technology change over the past five years.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Factors which care providers' identified as greatly influencing their acceptance of technology included ease of use; providing continuity of patient care; and improving quality of patient care. When asked to identify whether technology had the potential to provide benefit throughout a number of identified areas, the overwhelming response was that technology has great potential to provide benefits, however to date has not fully realised its actual potential.

When respondents were asked about Commonwealth and State legislation relating to recordkeeping, and the subsequent management of patient information, approximately three-quarters of respondents (73.6%) indicated that they did not have an understanding of the relevant Commonwealth or State legislation. Care providers also indicated that increased privacy legislation has had a considerable impact on their work practices. Increased privacy legislation has meant that care providers' have had to improve documentation relating to patient information, while it has also made it more difficult to access records and contact relatives of patients.

Specifically relating to SEHRs, over half of the respondents (57.6%) were not aware of trials which have been conducted throughout Australia. However, of those aware of the trials, the majority indicated that the results of these trials would not influence their willingness to adopt SEHRs (82.6%). Generally, respondents were more willing to adopt SEHRs in the future (72%) rather than currently (63.5%). The factors identified which would most positively influence care providers' willingness to adopt SEHRs related to improved patient care, reduced workloads and reduced complexity. The main problems identified by care providers relating to implementation of SEHRs included privacy/complexity and compliance issues.

Overall, based on an assessment of general attitudes towards technology and SEHRs, the majority of respondents could be assumed as having a positive attitude towards the use of technology for general day to day use, and also for patient care. However, care providers indicated that the expected benefits of technology have not typically been realised to date. Shared electronic health records divided opinion, with the majority of respondents indicating some level of degree of willingness to adopt and implement SEHRs, provided various potential benefits were realised. Based on care provider perceptions of a SEHR system such as HealthConnect/PCEHR, the biggest challenge to overcome would appear to be the critical issues of complexity and factors relating to relative advantage such as workloads, and demonstrating the ability of SEHRs to achieve stated benefits.

Findings revealed that Relative advantage had a statistically significant potential to influence care providers' future willingness to adopt SEHRs. Additionally, dimension one (perceived benefits of technology) was found to have a greater potential to influence care providers' future willingness to adopt SEHRs compared to the variables comprising dimension two (perceived advantages of SEHRs) of Relative advantage.

Multiple regression analysis for Compatibility found that this factor had the potential to have a statistically significant influence on care providers' future willingness to adopt SEHRs. Of the two dimensions comprising Compatibility, dimension two (perceived compatibility of SEHRs with care providers' values relating to patient care) had a greater potential to influence willingness to adopt compared to the variables best described as perceived potential to improve communication between care providers (dimension one).

Multiple regression analysis found that Complexity also had the potential to have a statistically significant influence on care providers' future willingness to adopt SEHRs. Of the two dimensions comprising Complexity (perceptions of complexity relating to information and records management considerations and perceptions of the complexity relating to compliance), only dimension two produced statistically significant results.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Following the testing of hypotheses one to three, the data was then analysed to determine whether any variation existed between the three tiers of the health system within Queensland future willingness to adopt SEHRs (i.e. GPs, Specialists and Hospital staff). Findings revealed that there was a statistically significant variation at the $\alpha = 0.05$ level in future willingness to adopt SEHRs across the three tiers, therefore supporting null hypothesis four. Responses indicate that overall, Hospital staff were more willing to adopt SEHRs with GP's the least willing based on a comparison of responses to 'Extremely willing' and 'Very willing'.

Finally, variation to future willingness to adopt SEHRs between the public and private sector was investigated in hypothesis five. The data revealed that there was sufficient evidence to support the null hypotheses at the $\alpha = 0.05$ level, indicating that a statistically significant variation existed between the public and private sector's future willingness to adopt SEHRs. While both the public and private sector featured similar response rates to 'Extremely willing' or 'Very willing' to adopt SEHRs in the future, the private sector contained a much greater number or respondents who were not willing to adopt (with responses to 'Extremely against' or 'Some concerns') compared to similar responses from public sector respondents.

Following the data analysis for the study's hypotheses, potential limitations were then addressed. Potential limitations were identified in two key areas, (a) the timing of the survey and (b) statistical testing methods. The timing of the study was identified as a potential limitation because it was likely that the respondents to the study had a limited knowledge of SEHRs (due to limited trials or implementation throughout Queensland, and Australia), which meant that the results did not correspond with findings from the literature review. The statistical testing methods posed potential limitations as the impact of Relative advantage, Compatibility and Complexity on willingness to adopt was not clear. Subsequently, logistic regression was used to identify whether the strength of the findings could be improved. Results of the logistic regression were consistent with previous testing (i.e. that Relative advantage, Compatibility and Complexity had the potential to be statistically significant predictors of care providers' future willingness to adopt SEHRs), and provided improved results relating to the strength of the relationships.

ACKNOWLEDGMENT

The research is part of the PhD work conducted by the first author (Dr Pullen) under the supervision of Dr H. Maguire and the second author, Associate Professor L. Al-Hakim. The PhD degree was granted to Dr Pullen by the University of Southern Queensland in 2012. The authors would like to thanks the Australian Medical association of Queensland (AMAQ) for the valuable support and assistance in distributing the questionnaire among care providers in Queensland. The authors would like to extend their thanks to Queensland State Archives for their support.

REFERENCES

Aap. (2004). *NSW medical records go online*. Retrieved 13/1/05, from http://news.ninemsn.com.au/article.aspx?id=8111

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Berner, E., Detmer, D., & Simborg, D. (2005). Will the wave finally break: a brief view of the adoption of electronic medical records in the United States. American Medical Informatics Association.

Briggs, L. (2000). *A Health Information Network for Australia*. Report to Health Ministers by the National Electronic Health Records Taskforce.

Burgess, M., Gray, W., & Fiddian, N. (2004). *Quality Measures and the Information Consumer.* Paper presented at the 9th International Conference on Information Quality, Cambridge, MA.

Burke, D., & Menachemi, N. (2004). Opening the Black Box: Measuring Hospital Information Technology Capability. *Health Care Management Review*, *29*(3), 207–217. doi:10.1097/00004010-200407000-00006 PMID:15357231

Burke, L., & Weill, B. (2005). *Information Technology for the Health Professions* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Claburn, T. (2004). *Software Aggregates Medical Knowledge in Hopes of Averting Misdiagnosis*. Retrieved 28/4/05, from http://www.informationweek.com/story/showArticle.jhtml?articleID=56200189

Clarke, R. (2001). *Research Challenges in Emergent e-Health Technologies*. Retrieved 13/1/05, from http://www.anu.edu.au/people/Roger.Clarke/EC/eHlthRes.html

Commonwealth of Australia. (2011). *Concept of Operations, Personally Controlled Electronic Health Record (PCEHR)*. Retrieved 26 May 2015, from Commonwealth Department of Health and Ageing http://ehealth.gov.au/internet/ehealth/publishing.nsf/content/CA2579B40081777ECA2578F800194110/$File/PCEHR-Concept-of-Operations-1-0-5.pdf

Conrick, M. (2006). *Health Informatics Appendix A What is Health Informatics*. Sydney: Thomson.

Darr, A., Harrison, M., Chakked, L., & Shalom, N. (2003). Physicians' and nurses' reaction to electronic medical records. *Journal of Health Organization and Management*, *17*(5), 349–359. doi:10.1108/14777260310505129 PMID:14628488

Dearne, K. (2005). $18m central e-health body. *The Australian*. Retrieved from http://www.theaustralian.news.com.au/common/story_page/0,5744,12106623%255E23289,00.html

Detmer, D., & Gillings, D. (2000). Information technology for quality health care: A summary of United Kingdom and United States experiences. *Quality in Health Care*, 9. PMID:10980079

Farman, D., Honeyman, A., & Kinirons, M. (2003). Risk reduction in general practice: The impacts of technology. *International Journal of Health Care Quality Assurance*, *16*(5), 220–228. doi:10.1108/09526860310486660

Furukawa, M., Raghu, T., Spaulding, T., & Vinze, A. (2008). Adoption of health information technology for medication safety in U.S. Hospitals, 2006. *Health Affairs*, *27*(3), 865–875. doi:10.1377/hlthaff.27.3.865 PMID:18474981

Goldstein, G., & Hersen, M. (1984). *Handbook of Psychological Assessment*. New York: Pergamon Press.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1998). *Multivariate Data Analysis* (5th ed.). Sydney: Prentice Hall International Inc.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hare, M. (2001, July). Information Technology in General Practice: Benefits & Barriers. *NIS Journal Watch*.

Health Canada. (2004). *Towards an evaluation framework for electronic health records initiatives: A review and assessment of methods used to measure the impact of health information systems projects*. Retrieved 28/2/05, from http://www.hc-sc.gc.ca/ohih-bsi/pubs/kdec/nf_eval_rpt2_e.html

Health Connect. (2004a). *Fact Sheet - HealthConnect*. Retrieved 13/1/05, from http://www.healthconnect.gov.au/pdf/FactSheet_0409.pdf

Health Connect. (2004b). *HealthConnect - an overview*. Retrieved 14/1/05, from http://www.healthconnect.gov.au/pdf/overviewDec04.pdf

Jaccard, J., & Wan, C. (1996). *LISREL approaches to interaction effects in multiple regression.* Thousand Oaks, CA: Sage Publications.

Jefferson, J. (1998). A study into levels of, and attitudes towards information technology in general practice. Retrieved 28/2/05, from http://www7.health.gov.au/pubs/gpit/

Johnson, J. (n.d.). *Visual Analog Scales: Part III*. Retrieved 3/01/2004, from www.bamc.amedd.army.mil/DCI/articles/dci079676.htm

Keeley, D. (2000). Information for health - hurry slowly [editorial]. *The British Journal of General Practice*, *30*(4).

Kerr, K., & Norris, T. (2004). *The Development of a Healthcare Data Quality Framework and Strategy.* Paper presented at the 9th International Conference on Information Quality, Cambridge, MA.

Kohn, L. (2000). *To Err is Human: Building a Safer Health System*. Washington, DC: National Academy Press.

Lorence, D., & Jameson, R. (2002). Adoption of information quality management practices in US healthcare organisations. *International Journal of Quality & Reliability Management*, *19*(6), 737–756. doi:10.1108/02656710210429591

Lott, L. (1997). The conundrum of patient records. *Leadership in Health Services*, *6*(2). PMID:10169228

McGee, M. (2004). *Execs Support E-Health Records*. Retrieved 10/1/2005, from http://nwc.complliancepipeline.com/news/52500182

Merkes, M. (2000). Electronic Health Records - Your Health Online. *PC Update Online, 17*(8).

Mitchell, E., & Sullivan, F. (2001). A descriptive feast but an evaluative famine: Systematic review of published articles on primary care computing during 1980-97. *BMJ (Clinical Research Ed.)*, 320. PMID:11157532

More, E., & McGrath, M. (2002). An Australian case in e-health communication and change. *Journal of Management Development*, *21*(8), 621–632. doi:10.1108/02621710210437590

Mozian, J. (2003). Beyond the biz. *B to B, 88*(5).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Murray, M., & Lynn, M. (1996). Developing microcomputer-based quality assurance systems for health care. *International Journal of Health Care Quality Assurance*, *9*(6), 10–17. doi:10.1108/09526869610128160 PMID:10162124

Naeme, R., & Kluge, E. (1999). Computerisation and health care: Some worries behind the promises. *BMJ (Clinical Research Ed.)*, 319.

Nicastro, E. (1999). Cover letter should yield an immediate response. *Puget Sound Business Journal, 20*(25).

Richards, B., Bolton, P., Veale, B., & Quinlan, F. (1999). *Information technology in general practice*. Canberra: Commonwealth Department of Health and Aged Care.

Royal Australian College of General Practitioners. (2004). *A Quality Framework for Australian General Practice: Background Paper*. Author.

Schloeffel, P. (2004a). *Current EHR Developments: an Australian and International Perspective - Part 1*. Retrieved 28/2/05, from http://www.enigma.co.nz/hcro/website/index.cfm?fuseaction=articledisplay&FeatureID=010904

Schloeffel, P. (2004b). *ISO/DTR 20514. Health informatics - electronic health record: definition, scope and context*. Fourth Draft.

Shaughnessy, J., & Zechmeister, E. (1990). *Research Methods in Psychology* (2nd ed.). Sydney: McGraw-Hill.

Sisson, D., & Stocker, H. (1989). Analyzing and interpreting Likert-type survey data. *Delta Pi Epsilon Journal*, *31*(2).

Tornatzky, L., & Klein, K. (1982). Innovation charecteristics and innovation adoption implementation: A meta-analysis of findings. *IEEE Transactions on Engineering Management*, 29.

Zikmund, W. (2003). *Business Research Methods* (7th ed.). Australia: Thomson.

KEY TERMS AND DEFINITIONS

Compatibility: Compatibility can be viewed in terms of being consistent with existing values and procedures. In relation to innovation, compatibility is the degree to which an innovation is perceived as being consistent with the existing values, experiences, and needs of potential adopters.

Complexity: In relation to innovation, complexity is the degree to which an innovation is perceived as difficult to understand and use. New ideas that are simpler to understand will be adopted more rapidly than innovations that require the development of new skills and understandings.

Diffusion: Diffusion is an action of spreading something widely. In this chapter, this term is associated with innovation.

E-Health: E-health is a reference to the process of managing health information electronically.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Electronic Health Record: A repository of information regarding the health of a subject of care (individuals) in a computerised form, stored and transmitted securely, and accessible by multiple authorised users such as general practitioners and emergency doctors.

Healthcare Provider: An individual provides healthcare services to patients.

HealthConnect: HealthConnect is Australia's first health information network, which involves the collection, storage and exchange of individual's health information via a secure network.

Observability: Observability relates to those involved in the adoption process being able to observe the results of the adoption. In relation to innovation, observability can be defined as the degree to which an innovation are visible to others. More observable innovations are more likely to be adopted.

Relative Advantage: In relation to innovation, relative advantage gauges the degree of advantage gained from implementation of an innovation, compared to the previous systems used. It can also be described as the degree to which an innovation is perceived as better than the idea it supersedes. The greater the perceived relative advantage of an innovation, the more rapid will be its rate of adoption.

Trialability: The notion of trialability refers to the ability to trial an idea, process or system before making the eventual decision to implement or not to implement the idea, process or system. In relation to innovation, triability can be described as the degree to which an innovation may be experimented with on a limited basis.

This research was previously published in the Handbook of Research on Driving Competitive Advantage through Sustainable, Lean, and Disruptive Innovation edited by Latif Al-Hakim, Xiaobo Wu, Andy Koronios, and Yongyi Shou; pages 500-533; copyright year 2016 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 72
Electronic Medical Record Implementation Challenges for the National Health System in Greece

Dimitrios G. Katehakis
Institute of Computer Science (ICS), Greece

ABSTRACT

The purpose of this work is to expose challenges related to the implementation of quality electronic medical record (EMR) systems in public hospitals in Greece, a country where the national health system (NHS) has already acquired electronic medical records (EMRs). The level of EMR implementation, together with organizational maturity at a hospital level, are explored. What is discovered is that there are different adoption levels, not recorded in a systematic manner. The majority of physicians are either reluctant to implement EMRs or do not know options available to them. Implications include not continuous flow of events, cut off of critical information, lower quality of health services, patients not empowered to carry with them clinically significant information, unnecessary repetition of medical procedures and higher costs. It is concluded that focus should be paid on enabling the use of quality, interoperable and secure EMRs to better support medical decision, in an effort to improve the health of the population and to better control costs.

INTRODUCTION

Hospitals are complex institutions, devoted to the provision of patient treatment. They have specialized administrative, medical and nursing staff, as well as medical equipment. A hospital information system (HIS) encompasses the set of information technology (IT) applications used to manage hospital operations concerned with the administrative and clinical aspects of providing medical services within its boundaries. Subsequently, clinical and medical patient care activities are considered to be at the core of any HIS.

DOI: 10.4018/978-1-7998-1204-3.ch072

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The EMR is defined as the electronic patient record created by providers for specific encounters in hospitals and ambulatory environments (Habib 2010; Kierkegaard, 2011). In contrast, the electronic health record (EHR) is a longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting (Health Information and Management Systems Society, 2011). Unlike the EMR and the EHR, the personal health record (PHR) refers to a representation of health records related to the care of the patient that is managed by the patient (Tang et al., 2006).

With the growth of mobile computing, the number of records regarding personal health is increasing exponentially. That leads to the need for an integrated method of storing health-related data which could be used by health care providers and patients (Roehrs et al., 2017). Connecting to the EMR requires shifting from paper, establishing interfaces to link islands of information based on certain workflows, and semantically homogenising the produced information. It is not difficult for any health professional to see the direct benefits of using EHR and having both administrative and clinical data that are accessible, comparable, communicable, and confidential (Iakovidis, 1998). Streamlining patient information flow and its accessibility to other healthcare providers through national, regional and/ or patient controlled services has the potential to improve care quality and patient safety over time.

Today, comparable and high-quality administrative and clinical data form the basis for effective and efficient management of public and clinical health. All current national initiatives focus on how to control costs and improve services. That is why electronic prescription (ePrescription – the ability to send readable and error-free prescriptions electronically from the healthcare provider to the pharmacy) and electronic reimbursement (eReimbursement – the ability to submit and follow-up electronically on claims with health insurance organizations) services seem to be of high priority together with electronic identification initiatives. Development of national infrastructures to support cross-border services is high on the agenda (World Health Organization and International Telecommunication Union, 2012; European Commission, 2012) and a move towards connecting the patient's team of healthcare providers to facilitate informed decision making.

Healthcare in Greece is provided by the NHS, or ESY (Greek: Εθνικό Σύστημα Υγείας, ΕΣΥ). It consists of a universal health care system provided through national health insurance, and private health care. In 2000, the Greek government, in accordance with the European Union (EU) recommendations, procured an operational program for implementing the information society strategy for Greece in a coherent and integrated way (Katehakis et al., 2011). The program aimed to provide IT support for regional health authorities and the public health organizations they supervise, to achieve better management of human resources, performance, containment of total cost, and continuous improvement of processes. Foreseen benefits included an upgraded quality of services to citizens through business process re-engineering and reduction of medical errors, secure exchange of medical information, and efficient access to the EHR.

That resulted in the introduction of modern IT systems for the vast majority of hospitals in the country. Incidentally, the conclusion of this effort coincided with the beginning of the economic crisis in the country, which subsequently led to an increasing number of admissions to public healthcare facilities, and deterioration of self-rated health (Kastanioti et al., 2013; Simou et al., 2014). The intention of the paper is not to discuss the challenges of the Greek healthcare system but rather challenges related to EMR implementation within the Greek NHS. An in-depth description of the NHS in Greece is beyond the scope of this work.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In the section that follows the methods used for conducting this work are described, EMR characteristics are outlined, the digitization status in the NHS of Greece is presented, EMR adoption at a hospital level is explored together with key factors contributing to EMR adoption. In the results section, EMR implementation challenges in Greece are described, followed by some discussion and conclusions in the corresponding sections.

METHODS

In order to conduct this work, a literature review was performed on established work, and information was retrieved from available published material and websites. Emphasis was paid on examining available source in regards to clarifying EMR characteristics, documenting EMR implementation status in Greece, reviewing work related to EMR adoption levels, and examining factors affecting EMR and EHR adoption.

In the following subsections, EMR characteristics are clarified, and the distinction between the concepts of EMR, EHR, and PHR is further explained. EMR implementation status in Greece, relevant legislation, and plans for the future are presented. Reference is made to the extent that medical files are digitized and interconnected as well as to remaining steps that have to be taken so that full advantage of interoperability is achieved on the road towards the EHR and efficient linking with personal health records (PHRs). The Healthcare Information and Management Systems Society (HIMSS) approach towards establishing EMR adoption levels is examined as well as factors contributing to EMR and EHR adoption.

EMR Characteristics

The EMR consists of electronic copies of medical records produced in a physician's office, a clinic or a hospital. It can include patient, hospital, and healthcare professional details, encounter details, clinical information, diagnoses, medication, operations, treatments, etc. The most important quality EMR systems characteristic is related to access to information in a secure and legitimate manner.

EMRs are meant to create legible and organized recordings and to access clinical information about individual patients and can serve as a data source for an EHR. The EHR is an essential tool for improving both the safety and quality of health care. However, physicians must actively use these systems to accrue the benefits (Ajami et al., 2013). The gains from EHR systems depend on meaningful sharing and rely on access to information regardless of place and time, and from re-using information for multiple purposes (Dobrev et al., 2010; Prey et al., 2016). That results in less time searching for information and more quality time with patients. Improved organisational efficiency and secondary uses of data are typically amongst the most commonly expected benefits (Black 2011). The EHR usually reflects the partial view of a health care provider without the ability for patients to control or interact with their data.

The EMR can be used to compose the EHR and is patient centric. EMRs and electronic health records (EHRs) are used by physicians to improve the quality of care and contain costs. Whereas EMR is usually considered an internal organizational system, the EHR is defined as an inter-organizational system. The EHR is primarily episodic, could be longitudinal, and incorporates administrative, financial, and clinical data. The concept of EHR extends beyond the boundaries of a health care provider and includes the concept of timelessness. Main advantage of EHR is the availability of cross-provider medical information.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The use of EHR requires the presence of a certain user and system attributes, and numerous organizational and environmental facilitators. That is why the EMR in a public hospital cannot be easily shared today among several hospitals. In many cases, this information exchange cannot be made electronically even within the boundaries of a single hospital.

Healthcare providers play a critical role in the adoption and utilization of health IT, including EMRs in hospitals (Cresswell et al., 2015). However, many healthcare professionals, from nurses to ambulatory care physicians, are still reluctant to utilize the available technologies, leading to limited system use and, eventually, system failure (Bowens et al., 2010; Linder, 2006).

Recently, a digital platform for patient-centred medical care known as PHR was introduced, as an enabler for self-management of medical records. PHRs are online systems used by patients, supporting a person-centred functionality, and could be cradle-to-grave. They support the shift from institution-centred care to citizen-centred care, paying emphasis on continuity of care from prevention to management of chronic conditions. On the contrary to EMRs and EHRs, the PHR is not a legal record. PHR enables integration of main information components in the EMR and EHR systems.

EMRs, EHRs, and PHRs can reside on different platforms under various technologies and standards. Although EMRs contain local information and provide fast and accurate delivery, the major advantage of EHR in medical practice is the availability of cross-provider medical information. Medical information integration leads to a dramatic change in personalized care but also poses serious challenges and threats to security and privacy. (Heart et al., 2016).

EMR Implementation Status in Greece

During the period 2000-2006, in the frame of the 3rd Community structural assistance programme, Greece started developing IT solutions for the majority of the –at that time- seventeen Health Regions. More than 60% of the NHS hospitals of Greece were covered by the programme and had purchased integrated information systems to support the regional health authorities they belonged to. These included:

- Enterprise resource planning (ERP) systems to support costing, general ledger, analytical accounting, budgeting, payable accounts, receivable accounts, human resources management, assets management, materials management, supply chain management, funds management, contract management, etc.;
- Patient administration applications for patient management such as admissions, discharge, transfers, bookings, accounts, etc.;
- Medical applications to support all healthcare organization processes related to the provision of medical and nursing care for outpatient clinics, emergency rooms, diagnostic procedures, medical documentation, doctor letters, medical history, order entry management, nursing stations, etc.;
- Laboratory information systems (LIS);
- Radiology information systems (RIS);
- In limited cases picture archive and communications systems (PACS);
- Management information systems (MIS);
- Applications for primary care to support the electronic storage and management of patient data, regarding visits to primary care providers.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

However, as the development of the "Integrated Information System" for each region was procured separately, they resulted in different systems without interoperability among them, as well as between hospitals belonging to the same health region in several cases. That was further complicated following the merging of the seventeen first health regions into seven, and the (at the moment) ongoing effort to introduce a new integrated information system to support hospitals not financed by the 3rd Community structural assistance programme. In addition, in 2011, the Ministry of Health and Social Solidarity announced its intention to cut back the number of beds and hospitals in the country from 131 hospitals with 35,000 beds to 83 hospitals with 33,000 beds. An attempt to have the archives of patient medical records digitized in 2014 was suspended in 2015.

The latest published information, in regards to hospital digitization status at the Hellenic NHS, is depicted in Table 1 (Hellenic Ministry of Health and Social Solidarity, 2012). It is important to note that almost all hospitals in the country were operating integrated hospital information systems in 2012; even partially. More than 70 hospitals reported operating an EMR for a basic package of medical, nursing and/ or administrative processes, as displayed in Table 1.

Today, despite the positive effects of EMR usage in medical practices, the level of digitization of EMRs in the Greek NHS is not recorded in a systematic manner and progress made is not reported accurately. No conclusion can be reached in regards to hospital EMR implementation maturity level in the NHS hospitals of Greece, apart from the fact that EMR adoption lags behind all other HIS subsystems. Paper still rules in most of the cases, especially as far as access to legacy records is concerned.

Greek eHealth policy (2014-2020) set as priorities the restructuring of primary healthcare, pooling of financial resources, introducing new managerial and administrative methods, adopting cost effectiveness and monitoring mechanisms, and developing policies for better resource allocation. The Ministry of Health (MoH) has the overall responsibility for electronic health (eHealth) in Greece. The development of an electronic patient record is a major objective and a priority of the National Health System (Milieu Ltd and Time.lex, 2013; EU Directorate-General for Health and Food Safety, 2014).

In November 2010 Law 3892/ 2010 put in place the national ePrescription system and the obligation to submit prescription and dispense medications electronically. The law describes duties of doctors and pharmacists, as well as access rights including for patient access to own information.

Table 1. Operation of basic information subsystems in hospitals of the NHS in Greece (as of 2012)

Information System	Full Operation	Partial Operation	Level of Adoption
Pharmacy	131	0	100%
Materials Management	130	1	99%
Patients Accounting Office	131	0	100%
Admissions Office	130	1	99%
Outpatient Clinics Admissions Office	108	11	91%
General Ledger	129	2	98%
Human Resources	69	15	58%
Ward Management – ePrescription	100	12	89%
LIS	81	20	80%
EMR	30	43	41%

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In December 2013 Law 4213/2013, transposing Directive 2011/24/EU (Quinn et al., 2011) in Greece on the application of patients' rights in cross-border healthcare and other provisions, was published.

In February 2014 Law 4238/2014 on the Primary National Health Network placed the responsibility for primary care under the auspices of the Ministry of Health. Following that law, Greece has set the legislative framework for the adoption of an EHR system. Article 51(4) provides for the creation of EHRs for all Greek citizens. It also provides for the inclusion of a minimum patient summary (PS) data set, as well as the responsibility of doctors for the creation and maintenance of EHRs.

During the past few years a multitude of eHealth services have been introduced to support ePrescription for primary care, eConfirmation for insurance status verification, eReimbursement (e-ΔΑΠΥ), electronic booking, and a business intelligence system that is in a position to automatically retrieve operational data to feed the MoH.

Priorities set in September 2016 by the National Council of Electronic Health Governance (Εθνικό Συμβούλιο Διακυβέρνησης Ηλεκτρονικής Υγείας) were the following:

- Creation of the National ehealth Interoperability Framework to support standardized interfaces;
- Creation of an eprescription services framework to support the functional integration of services nationwide;
- Creation of the legal framework for the operation of a national EHR;
- Creation of the National Contact Point for ehealth to support cross-border ehealth services.

At the time, this paper was prepared no national EHR system had been realised in Greece.

The HIMMS EMR Adoption Model

Understanding the performance of EMR in hospitals is a challenge in the health care context, and to that extent, several models have been developed so far (Carvalho, 2016). One of the most prominent models is the HIMSS Maturity Model for EMR used to identify EMR stages of maturity from the limited departmental systems to paperless EMR environment (HIMSS Analytics, 2008). The HIMSS EMR Adoption Model (EMRAM) consists of eight stages, described in Table 2. According to HIMSS Analytics, the structure of this model ensures that a stage is reached only when all their applications are operational. Every level is described in terms of its requirements.

HIMSS facilitates the development of IT solutions for healthcare mostly in order to systematically improve its quality, safety, and efficiency. The vision includes the creation of mutually interconnected and harmoniously co-functioning EMRs of patients which would connect distinct entities on the market and service providers.

EMRAM is a certification method that makes an assessment of maturity and degree of implementation of EMR possible and it has acquired worldwide recognition. The programme is a reference point for improving the level of implementation of IT solutions, and for moving away from medical documentation on paper to electronic medical documentation. If the percentage of health units belonging to each of these levels is monitored in a systematic manner, then it is possible to measure the progress being made on the road towards the full paperless hospital. That already happens in Canada and the US (Garets et al., 2006; HIMSS Analytics, 2008).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. The EMRAM reference system determines the level of EMR implementation in a healthcare organization (i.e. a hospital) in a scale from 0 to 7. Level 7 corresponds to a fully-paperless hospital (i.e. all patient records are kept in electronic form). That enables the managers of such establishments to see what level they are on and what steps they would have to take to improve their score.

Stage	Cumulative Capabilities
Stage 7: Complete EMR, Data Analytics To Improve Care	The hospital no longer uses paper charts to deliver and manage patient care and has a mixture of discrete data, document images, and medical images within its EMR environment. Data warehousing is being used to analyse patterns of clinical data to improve the quality of care and patient safety and care delivery efficiency. Clinical information can be readily shared via standardized electronic transactions (i.e. continuity of care document) with all entities that are authorised to treat the patient, or a health information exchange (i.e., other non-associated hospitals, ambulatory clinics, sub-acute environments, employers, payers and patients in a data sharing environment). The hospital demonstrates summary data continuity for all hospital services (e.g. inpatient, outpatient, emergency department, and with any owned or managed ambulatory clinics). Blood products and human milk are included in the closed-loop medication administration process.
Stage 6: Physician Documentation (Templates), Full Clinical Decision Support Systems (CDSS), Closed Loop Medication Administration	Full physician documentation with structured templates and discrete data is implemented for at least one inpatient care service area for progress notes, consult notes, discharge summaries or problem list & diagnosis list maintenance. Level three of clinical decision support guides all clinician activities related to protocols and outcomes in the form of variance and compliance alerts. The closed loop medication administration with the bar coded unit dose medications environment is fully implemented. The electronic medication administration record application (EMAR) and bar coding or other auto identification technology, such as radio frequency identification, are implemented and integrated with computerised practitioner order entry (CPOE) and pharmacy to maximise point of care patient safety processes for medication administration. The "five rights" of medication administration are verified at the bedside with the scanning of the bar code on the unit does medication and the patient identification (ID).
Stage 5: Full radiology PACS	A full complement of radiology PACS provides medical images to physicians via an intranet and displaces all film-based images. Cardiology PACS and document imaging are scored with extra points.
Stage 4: CPOE; Clinical Decision Support (Clinical Protocols)	CPOE for use by any clinician licensed to create orders is added to the nursing and clinical data repository (CDR) environment along with the second level of clinical decision support capabilities related to evidence-based medicine protocols. If one inpatient service area has implemented CPOE with physicians entering orders and completed the previous stages, then this stage has been achieved.
Stage 3: Clinical Documentation, CDSS (Error Checking)	Nursing/clinical documentation (e.g. vital signs, flow sheets, nursing notes, EMAR is required and is implemented and integrated with the CDR for at least one inpatient service in the hospital; care plan charting is scored with extra points. The EMAR is implemented. The first level of clinical decision support is implemented to conduct error checking with order entry (i.e., drug/drug, drug/ food, drug/lab conflict checking normally found in the pharmacy information system). Medical image access from PACS is available for access by physicians outside the Radiology Department via the organisation's intranet.
Stage 2: CDR, Controlled Medical Vocabulary, CDS, health information exchange (HIE) Capable	Major ancillary clinical systems feed data to a CDR that provides physician access for reviewing all orders and results. The CDR contains a controlled medical vocabulary and the clinical decision support/rules engine (CDS) for rudimentary conflict checking. Information from document imaging systems may be linked to the CDR at this stage. The hospital may be HIE capable at this stage and can share whatever information it has in the CDR with other patient care stakeholders.
Stage 1: All Three Ancillaries Installed - Lab, Rad, Pharmacy	All three major ancillary clinical systems are installed (i.e., pharmacy, laboratory, and radiology).
Stage 0: All Three Ancillaries Not Installed	The organization has not installed all of the three key ancillary department systems (laboratory, pharmacy, and radiology).

Understanding the level of medical records digitization within a health unit is a colossal challenge for Greek hospitals. Today, very few hospitals have EMR solutions that can effectively reduce medical errors or improve the quality and efficiency of patient care. The EHR can be established only if the EMRs of the various care delivery organizations have evolved to a level that can create and support a robust exchange of information between the involved stakeholders.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Factors Contributing to EMR and EHR Adoption

It is commonly agreed that information contained in the EHR is critical to support clinical care and the longitudinal EHR is of high value. Still, it must be collected and assembled from multiple sources, and this is particularly challenging. The assembly of quality, semantically meaningful and consistent data has been until now almost impossible, and this is the reason why so many efforts are made towards developing EHRs.

Since health information is regarded by many as being among the most confidential of all types of personal information (Fernández-Alemán, 2012), security concerns are considered to be one of the main reasons behind resistance shown to widespread EHR adoption (Boonstra et al., 2014). From the physicians' perspective, concerns regarding privacy and confidentiality have been found to hinder physicians' acceptance and EHR adoption, demonstrating how these issues can impact EHR use that extends beyond patients' reservations (Ben-Assuli, 2015).

In the past three decades, several studies have extracted antecedents to the user adoption of health information systems. According to (Najaftorkaman et al., 2014) factors contributing to the user adoption of EMRs are classified in eight categories: individual factors, psychological factors, behavioural factors, environmental factors, organizational factors, financial factors, legal factors, and technical factors. Training is also considered to be a critical part of EMR implementation and, as reported by (Nechyporenko et al., 2015), to perform it appropriately, an educational approach should be developed individually for each specific hospital facility.

However, knowing these factors do not necessarily lead to successful implementations unless the unique characteristics of the implementation environment are examined in relation to these factors, and particular attention is paid to encouraging and facilitating change. Implementation must be carefully coordinated and appropriately planned for the production of outcomes acceptable to the stakeholders involved. Organizational issues must be considered with caution, and the appropriate political support must be in place to decisively resolve outstanding issues.

RESULTS – EMR CHALLENGES IN GREECE

Although the majority of the hospitals of the NHS in Greece has acquired EMRs, the actual utilization of the investment made is still kept low. As already described there is no published evidence on how the EMR adoption level is progressing in Greece during the past few years. The latest published material comes from 2012.

While the EMR has many useful features to help optimize health care processes, some healthcare professionals find the whole system too complex and feel that they could provide excellent care to patients, even without an EMR. Others do not possess the financial resources to start efficient EMR use as they do not even know where and how to profile available systems to their individual needs. When applied correctly, the EMR serves as a very useful tool for maintaining critical medical information and increases the efficiency of workflows. That said, there are a number of challenges that must be overcome if one wants to reap the full benefits of EMRs.

Key challenges and concerns in Greece are related to linking with the workflow (lack of systems interoperability, this hinders successful adoption causing resistance to change), making best use of the champions (non-financial incentives are equally important to clinic staff as financial ones), and post

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

implementation follow up and support (financing the transition and sustainability model) to secure the way forward. The following sections focus on these factors that need to be taken into consideration towards effective exploitation of EMRs in the hospitals of the NHS in Greece.

Link with the Workflow

Usually, hospital infrastructures are closed networks restricted for administrative and clinical support. In order to support integrated workflows, it is vital that health IT systems adopt nationally accepted data standards that facilitate semantic interoperability. Besides the need for having established a national infrastructure, all relevant legal and regulatory enablers should need also be put in place as well.

Data entered once, used anywhere is the key concept. That requires fully interfaced ICT solutions in a stable and standardized manner. Fundamental principles to follow include the exploitation of available open standards and best practices, the adoption of an interoperability framework tailored to national needs, and implementation based on planning. This will allow for the development of quality EMRs, the development, maintenance and evolution of a scalable, secure, and affordable infrastructure to support incremental evolution building upon existing systems, while adding new capabilities as they become available. Issues related to usability and efficient linking with the workflow shall, therefore, need be addressed.

Today the means for establishing patient identification in Greece is based on AMKA, which is the social security number issued in the country. Common national identification approach should be followed in line with European Regulation (EU) No 910/2014, commonly known as eIDAS (Melin et al., 2016). In the future, a combination of national identity providers may provide the means towards proper end user authentication for EMRs.

Trust and security is a critical area in which healthcare presents demanding and challenging needs. IT security in healthcare systems, services, and applications is positioned as a major concern due to the high privacy and confidentiality requirements of sensitive healthcare data (ENISA, 2015). Private data needs to be accessed by professionals having a qualified role, with the consent of the patient and under strict audit controls. Medical data must only be disclosed or shared after the patient is identified and authenticated with sufficient accuracy.

In order to enable workflow automation for the effective sharing of clinically significant information EMR data will have to be shared with well-defined processes, EMRs must contain reliable content, and use agreed semantics, i.e. common terminology. That is expected to effectively facilitate, amongst other, organized referrals, e.g. from primary healthcare to hospital physicians.

Efforts are currently made towards that direction by the MoH for the creation of the National eHealth Interoperability Framework to support standardized interfaces which is expected to focus not only at a cross-border, national, or regional level but also within certain healthcare organizations (i.e. hospitals) and also to provide for reaching citizens at home, based on certain use cases. This ecosystem of services has the potential to tackle longstanding HIS challenges and open new avenues for research (Koumaditis et al., 2015). This is a strategic approach that needs to be further supported at all levels starting from technical, to semantic, to organisational, and all the way to legal. This effort should be aligned with available best practices, current experience, and available standards (eHealth Network 2015; European Commission, 2015; European Commission, 2017).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Make the Best Use of Champions

At the individual level, human factors play a major role in the design of an intervention, determining usability and ultimately adoption, despite the fact that organisational issues are critical in a strategically planning deployment that ultimately influences adoption (Greenhalgh et al., 2004).

Innovative technologies alone cannot lead to revolutionary changes. What is needed is cultural transformation and a new healthcare providers' cooperation model that does away with existing information silos. EHRs should not be looked as a side-process or administrative tool, but rather as an inseparable part of the health care process. Although it is the physician and the payer that frequently control the type and quality of care patient receives (Petru, 2016), implementing the smart hospital vision requires considering continuous adaptation, clinical outcome, efficiency in the supply chain and care processes and enhancement of own capabilities in regard to patient services (Kharbanda et al., 2017). Qualified human resources are a key ingredient for success. Education, training and continuous professional development for all, including for those citizens and patients which are capable and motivated to become engaged in their own care, should be strongly promoted (Stroetmann et al., 2011).

An important factor influencing user adoption is user confidence. Professionals should be encouraged with a clear motivation system and an explicit presentation of realistically envisioned benefits. Effective communication can facilitate engagement and acceptance. Individual initiatives should not be rejected, but if they meet certain quality criteria should be exploited to become part of the solution.

Finance the Transition and Sustainability Model

Quality and secure EMR deployment do not come without cost. Both initial costs, as well as the extra expenses to support personnel and operation of such systems, should be taken into consideration (Iakovidis, 1998). Without the right stimulation, hospitals will not be interested in improving their EMR-related competence (Hoyt, 2016). As already been suggested by (Fragidis et al., 2011), Greece needs to invest significant financial resources and work towards standards creation in order to gain the benefits of EHR for its citizens in the near future. Funding should be provided and assured, while responsibilities for sustainability should be clearly defined.

As far as Greece is concerned investments for initial acquisition and setup have already been made, as described in previous sections. Adoption level is still low, and no uniform way of evaluating and establishing comparable service level agreements with vendors providing technical support are in place. Without a considerable, coordinated effort, which will likely include realigning financial incentives, EHR adoption in hospitals will lag behind. That will have to be combined with the quality of care incentives in order for the EMR adoption to be successful. Nevertheless, these investments are unlikely to pay back unless end-users, healthcare providers and patients alike, endorse them. As already noted by (Berg, 2001), the success of any technological healthcare innovation will be eventually decided on the work floor.

DISCUSSION

Very often medical personnel in hospitals is called to support a clinical decision on top of missing or incomplete information. Diagnosis is mostly based on medical history, but medical histories may be incomplete in each EMR system and patient information dispersed in a manner that cannot be easily

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

assembled. Medical personnel spend significant time in search and re-entry of information. Today there is much physician dissatisfaction with EMRs due to poor usability, and it is not unreasonable to think that improving physician satisfaction with EMRs may also improve patient satisfaction (Kanter, 2016). Performing quality analysis on EMR implementation by capturing consumer expectations using multi-dimensional research instruments such as SERVQUAL (Parasuraman, 1988) may provide valuable insights in regard to the dimensions that are related to service quality gap when it is applied to different settings.

Ongoing work at an EU level seeks to create a common set of standards to embrace the national diversity and facilitate interoperability. Electronic prescribing is considered to be one of the major reasons for the push for EMRs. Large-scale deployment of cross-border ePrescription and PS services are expected to actively refine and improve already available software infrastructures in order to support patients' rights effectively in cross-border healthcare (Katehakis et al., 2017). In the case of PS, both patient and health professional may have access to the person's PS and other relevant documents as smoothly as at home. Top-down pressure call for integrated EMRs. On the other hand, EU member states need to be open and flexible in the design of joint solutions and to be able to ensure the implementation of the decided solutions nationally.

Patients are the ones who are expected to benefit most from the use of EHRs. For this to happen EMRs must be in a position to interoperate with each other. EMRs should be trustworthy, trusted, and usable for their purposes while individual patient health data should be reused. EHR is based on patient consent and requires security standards to create a sense of security and reliability. Implementations should follow modern approaches and provide technical measures that safeguard the privacy and confidentiality of the patients' EMRs, compatible with relevant data protection and confidentiality laws. The outcome should be acceptable to the medical community. The foreseen introduction of a new identity in Greece, encompassing digital features (like e.g. for digital signing), is expected to act as an enabler for safer EMRs.

The EC and the Member States together already explore cooperation opportunities in the wider eGovernment field on topics like ID management, secure infrastructures, data protection, access rights, thereby also avoiding the duplication of efforts while at the same time work on facilitators to boost evolution. The recently introduced General Data Protection Regulation enshrines a wide range of existing and new rights for individuals in respect of their personal data (The European Union, 2016). These include the right to be forgotten, the right to request the porting of one's personal data to a new organisation, the right to object to certain processing activities and also to decisions taken by automated processes. According to (ENISA, 2015) member states need to implement widely accepted security standards to achieve interoperability, while trustworthiness should be the responsibility of the country (i.e. national and/ or regional authorities). EHRs must protect both patients' EMRs and physicians' integrity.

Major challenges are likely to include continuing debates surrounding security issues (i.e. breaches of confidentiality that may lead to an erosion of confidence in the healthcare system, changing the way individuals interact with the system) and the capacity to meaningfully analyse data that become available (i.e. EMR data quality). According to (Accenture, 2016), for the majority of EHR systems, consumers believe that the benefits of sharing medical information electronically outweigh the risks. To them, effective treatment is the top priority. Social networking tools already enable patients to find each other, support each other, learn from each other and help each other to take decisions. All this happens outside of the healthcare system (Pitsch, 2016). From the perspective of patients, it is evident that solutions for sharing medical information electronically are needed.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

We will not be able to build the EHR unless the EMRs generated and maintained by physicians are put together. Key actions that need to be taken on the way towards quality EHRs are the following:

- Establish a national interoperability framework and let EMRs interoperate with each other and with EHRs. Widespread adoption by physicians will require the EHR system to make their professional lives easier, not more complex, and will need to provide a clear benefit to their clinical activities. To do this implementation should be meaningful and based on appropriate technical characteristics and common implementation standards to facilitate the improvement of data quality;
- Raise awareness. Expose best practices. Inform everyone that EMR systems are available. Meet expectations. Provide guidelines and resources for quality and secure enablement. Use incentives for motivating EMR adoption and use;
- Professionally establish robust eHealth services by financing standardization and quality assurance. Establish a way to monitor and control progress. A balance between long and short-term objectives should be maintained. The magnitude of the investment needed should not be underestimated if the EHR implementation is to be successful.

CONCLUSION

Our ability to connect places where healthcare is provided and sharing of medical data is expected to close the gap that traditionally limits the provision of high-quality health services. EMR is the cornerstone for providing content to evolving eHealth services. Yesterday this information was kept in a medical file closed in a doctor's cabinet, today it is processed by clinical decision support systems and tomorrow it will be analysed by artificial intelligence systems to support effective disease management and decision support. Semantic interoperability, when achieved, is expected to improve quality, safety, efficiency and efficacy of healthcare delivery. If we want to improve the health of the population, we need quality and secure EMRs.

The effort has to be paid towards giving meaning to the EMR systems already available at the Greek NHS. That will require gradual implementation based on planning as well as the evolution of the already existing infrastructure. The focus should be paid on operation and interconnection of individual EMRs as well as standardizing quality, in line with international practices. To this extent, effort will be required towards testing, certification, and application. At the same time, the development of comprehensive policies for privacy, consent management, and access to EMRs and the EHR will be required, together with the corresponding legislation and participation encouragement.

By building on existing human effort and experience and already available infrastructure, the significant funding that will be required can be controlled better. By having the entire EHR available, it will be far more possible to tailor medical care to each individual's clinical needs and empower healthcare professionals and citizens to take informed decisions. Furthermore, health IT will enable improved clinical outcomes, while at the same time benefit global research and science.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Accenture (2016). *Consumer Survey on Patient Engagement*. Retrieved March 27, 2017, from https://www.accenture.com/t20160629T045303__w__/us-en/_acnmedia/PDF-6/Accenture-Patients-Want-A-Heavy-Dose-of-Digital-Research.pdf

Ajami, S., & Bagheri-Tadi, T. (2013). Barriers for Adopting Electronic Health Records (EHRs) by Physicians. *Acta Informatica Medica*, *21*(2), 129–134. RetrievedMarch82017 doi:10.5455/aim.2013.21.129-134 PMID:24058254

HIMSS Analytics. (2008). Electronic Medical Record Adoption Model. Retrieved March 26, 2017, from http://www.himssanalytics.org/emram

Ben-Assuli, O. (2015). Electronic health records, adoption, quality of care, legal and privacy issues and their implementation in emergency departments. *Health Policy (Amsterdam)*, *119*(3), 287–297. Retrieved May 17, 2017 from https://doi.org/10.1016/j.healthpol.2014.11.014 doi:10.1016/j.healthpol.2014.11.014 PMID:25483873

Berg, M. (2001). Implementing information Systems in health care organizations: Myths and challenges. *International Journal of Medical Informatics*, *64*(2-3), 143–156. doi:10.1016/S1386-5056(01)00200-3 PMID:11734382

Black, A. D. Car1 J., Pagliari C., Anandan C., Cresswell K., Bokun T., McKinstry B., Procter R., Majeed A., & Sheikh A. (2011). The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview. *PLOS Medicine, 8*(1). doi: Retrieved January 4, 2017, from http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000387 doi:10.1371/journal.pmed.1000387

Boonstra, A., Versluis, A., & Vos, J. F. J. (2014). Implementing electronic health records in hospitals: A systematic literature review. *BMC Health Services Research*, *14*(1), 370. Retrieved May 17, 2017 from https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-14-370 doi:10.1186/1472-6963-14-370 PMID:25190184

Bowens, F. M., Frye, P. A., & Jones, W. A. (2010). Health Information Technology: Integration of Clinical Workflow into Meaningful Use of Electronic Health Records. *Perspectives in Health Information Management*, *7*(Fall). Retrieved May 19, 2017 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2966355/ PMID:21063545

Carvalho, J. V., Rocha, Á., & Abreu, A. (2016). Maturity Models of Healthcare Information Systems and Technologies: A Literature Review. *Journal of Medical Systems*, *40*(6), 131. Retrieved March 26, 2017 from https://link.springer.com/article/10.1007/s10916-016-0486-5 doi:10.100710916-016-0486-5 PMID:27083575

Cresswell, K. M., & Sheikh, A. (2015). Health information technology in hospitals: Current issues and future trends. *Future Hospital Journal*, *2*(1), 50–56.

Dobrev, A., Jones, T., Stroetmann, V., Stroetmann, K., Vatter, Y., & Peng, K. (2010). Interoperable eHealth is Worth it: Securing Benefits from Electronic Health Records and e-prescribing. European Commission, *Information Society and Media Study Report*, Bonn/ Brussels. Retrieved March 19, 2016, from http://www.ifap.ru/library/book476.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

eHealth Network (2013). *Guidelines on minimum/ nonexhaustive patient summary dataset for electronic exchange in accordance with the cross-border Directive 2011/ 24/ EU*. European Commission, Version 1.0. Retrieved March 19, 2016, from http://ec.europa.eu/health/ehealth/docs/guidelines_patient_summary_en.pdf

eHealth Network (2015). *Refined eHealth European Interoperability Framework*. Retrieved April 10, 2017, from http://ec.europa.eu/health//sites/health/files/ehealth/docs/ev_20151123_co03_en.pdf

ENISA. (2015). Security and Resilience in eHealth. Security Challenges and Risks. *European Union Agency for Network and Information Security*. Retrieved March 4, 2017 from https://www.enisa.europa.eu/publications/security-and-resilience-in-ehealth-infrastructures-and-services

EU Directorate-General for Health and Food Safety. (2014). *Overview of the national laws on electronic health records in the EU Member States and their interaction with the provision of cross-border eHealth services*, Brussels. Retrieved March 19, 2016, from http://ec.europa.eu/health/ehealth/docs/laws_report_recommendations_en.pdf

European Commission. (2012). *eHealth Action Plan 2012-2020: Innovative healthcare for the 21st century*. Retrieved March 4, 2017, from http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012DC0736

European Commission. (2015). *Commission Decision (EU) 2015/1302 of 28 July 2015 on the identification of 'Integrating the Healthcare Enterprise' profiles for referencing in public procurement (Text with EEA relevance)*. Retrieved April 8, 2017, from http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32015D1302

European Commission. (2017). *New European Interoperability Framework*. doi: Retrieved April 10, 2017 from https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf doi:10.2799/78681

European Union (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Retrieved March 13, 2017, from http://eur-lex.europa.eu/eli/reg/2016/679/oj

Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. A. O., & Toval, A. (2013). Security and privacy in electronic health records: A systematic literature review. *Journal of Biomedical Informatics*, *46*(3), 541–562. Retrieved May 17, 2017 from https://doi.org/10.1016/j.jbi.2012.12.003 doi:10.1016/j.jbi.2012.12.003 PMID:23305810

Fragidis, L. L., & Chatzoglou, P. D. (2011). The use of Electronic Health Record in Greece: Current Status. In *Proceedings of the 11th IEEE International Conference on Computer and Information Technology (CIT)* (pp. 475-480). 10.1109/CIT.2011.42

Garets, D., & Davis, M. (2006). Electronic Medical Records vs. Electronic Health Records: Yes, There Is a Difference. *A HIMSS AnalyticsTM White Paper*. Retrieved March 26, 2017, from http://s3.amazonaws.com/rdcms-himss/files/production/public/HIMSSorg/Content/files/WP_EMR_EHR.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Greenhalgh, T., Robert, G., Bate, P., Kyriakidou, O., Macfarlane, F., & Peackock, R. (2004). How to Spread Good Ideas: A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation. *Report for the National Co-ordinating centre for NHS Service Delivery and Organisation R & D (NCCSDO)*, 1-426. Retrieved March 27, 2017, from http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1201-038_V01.pdf

Habib, J. L. (2010). EHRs, meaningful use, and a model EMR. *Drug Benefit Trends*, *22*(4), 99–101.

Health Information and Management Systems Society. (2011). *Electronic Health Records*. Retrieved May 16, 2017 from http://www.himss.org/library/ehr

Heart, T., Ben-Assuli, O., & Shabtai, I. (2016). A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy. *Health Policy and Technology*, *6*(1), 20–25. Retrieved-March262017 doi:10.1016/j.hlpt.2016.08.002

Hellenic Ministry of Health and Social Solidarity. (2012). *Report of Hellenic Ministry of Health and Social Solidarity and National Health System Unit Results for 2011*, Athens, March 2012 (in Greek). Retrieved March 13, 2017, from http://www.moh.gov.gr/articles/hlektronikes-efarmoges-e-s-y/1332-ethsia-ekthesh-2011?fdl=4415

Hoyt, J. P. (2016). Record-breaking electronic efficiency. Interview included in "eHealth Trends & Talks". *Polish Healthcare Journal*, September 15, 69-72. Retrieved March 13, 2017, from https://www.osoz.pl/static_files/osoz/eHealth_2016.pdf

Iakovidis, I. (1998). Towards personal health record: Current situation, obstacles and trends in implementation of electronic healthcare record in Europe. *International Journal of Medical Informatics*, *52*(1–3), 105–115. Retrieved February 26, 2017 from http://www.sciencedirect.com/science/article/pii/S1386505698001294 doi:10.1016/S1386-5056(98)00129-4 PMID:9848407

Kanter, M. H. (2016). Capsule Commentary on Alkureishi et al., Impact of Electronic Medical Record Use on the Patient–Doctor Relationship and Communication: A Systematic Review. *Journal of General Internal Medicine*, *31*(5), 538. Retrieved March 28, 2017. doi:10.100711606-016-3626-1 PMID:26902240

Kastanioti, C., Kontodimopoulos, N., Stasinopoulos, D., Kapetaneas, N., & Polyzos, N. (2013). Public procurement of health technologies in Greece in an era of economic crisis. *Health Policy (Amsterdam)*, *109*(1), 7–13. Retrieved May 13, 2017. doi:10.1016/j.healthpol.2012.03.015 PMID:22502936

Katehakis, D. G., Halkiotis, S., & Kouroubali, A. (2011). Materialization of Regional Health Information Networks in Greece: Electronic Health Record Barriers & Enablers. *Journal of Healthcare Engineering*, *2*(3), 389–403. Retrieved December 30, 2016. doi:10.1260/2040-2295.2.3.389

Katehakis, D. G., Pangalos, G., & Prentza, A. (2017). Security Improvements for Better and Safer Cross-Border ePrescription and Patient Summary Services. *International Journal of Reliable and Quality E-Healthcare*, *6*(1), 18–28. doi:10.4018/IJRQEH.2017010102

Kharbanda, V. Car1 J., Bohlin, N., Sehlstedt, U., & Treutiger, J. (2017). *Building the Smart Hospital Agenda*. Retrieved February 1, 2017, from http://www.adlittle.com/downloads/tx_adlreports/ADL_Smart_Hospital.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kierkegaard, P. (2011). Electronic health record: Wiring Europes healthcare. *Computer Law & Security Report*, *27*(5), 503–515. doi:10.1016/j.clsr.2011.07.013

Koumaditis, K., & Themistocleous, M. (2015). Proposing and Testing SOA Governance Process: A Case Study Approach. *International Journal of Healthcare Information Systems and Informatics*, *10*(3), 42–62. doi:10.4018/IJHISI.2015070103

Linder, J. A. (2006). Barriers to electronic health record use during patient visits. *AMIA ... Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium*, *2006*, 499–503. Retrieved May 19, 2017 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839290/ PMID:17238391

Melin, U., Axelsson, K., & Söderström, F. (2016). Managing the development of e-ID in a public e-service context – challenges and path dependencies from a life-cycle perspective. *Transforming Government: People*, *Process and Policy*, *10*(1), 77–98. Retrieved May 19, 2017. doi:10.1108/TG-11-2013-0046

Milieu Ltd and Time.lex. (2013). Overview of the national laws on electronic health records in the EU Member States, National Report for Greece, *Report prepared by Milieu Ltd and Time.lex under Contract 2013 63 02*. Retrieved March 4, 2017, from http://ec.europa.eu/health//sites/health/files/ehealth/docs/laws_greece_en.pdf

Najaftorkaman, M., Ghapanchi, A. H., Talaei-Khoei, A., & Ray, P. (2014). A taxonomy of antecedents to user adoption of health information systems: A synthesis of thirty years of research. *Journal of the Association for Information Science and Technology*, *66*(3), 576–598. Retrieved March 4 2017 doi:10.1002/asi.23181

Nechyporenko, T., & McKibbon, A. (2015, October). Review of different training approaches to facilitate successful adoption of EMR systems by physicians in hospital settings (no. 54). *McMaster eBusiness Research Centre Working Paper*. Retrieved March 5, 2017, from https://macsphere.mcmaster.ca/bitstream/11375/20846/1/merc_wp_54.pdf

Parasuraman, A., Ziethaml, V., & Berry, L. L. (1988). SERVQUAL: A Multiple- Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, *62*(1), 12–40.

Petru, R. (2016). How to use EHRs to improve the quality of care? Interview included in "eHealth Trends & Talks". *Polish Healthcare Journal*, September 15, 69-72. Retrieved March 13, 2017, from https://www.osoz.pl/static_files/osoz/eHealth_2016.pdf

Pitsch, M. (2016). Dormant digitization. Interview included in "eHealth Trends & Talks". *Polish Healthcare Journal*, September 15, 62-64. Retrieved March 13, 2017, from https://www.osoz.pl/static_files/osoz/eHealth_2016.pdf

Prey, J. E., Polubriaginof, F., Kuperman, G. J., Tiase, V., Collins, S. A., & Vawdrey, D. K. (2016). International perspectives on sharing clinical data with patients. *International Journal of Medical Informatics*, *86*, 135–141. Retrieved March 21, 2016 from http://www.sciencedirect.com/science/article/pii/S1386505615300599 doi:10.1016/j.ijmedinf.2015.11.007 PMID:26631956

Quinn, P., & De Hert, P. (2011). The Patient's Rights Directive (2011/24/EU) – Providing (some) rights to EU residents seeking healthcare in other Member States. *Computer Law & Security Report*, *27*(5), 497–502. doi:10.1016/j.clsr.2011.07.010

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Roehrs, A., da Costa, C. A., da Rosa Righi, R., & de Oliveira, K. S. F. (2017). Personal Health Records: A Systematic Literature Review. *Journal of Medical Internet Research*, *19*(1), e13. Retrieved February 26, 2017 from http://www.jmir.org/2017/1/e13 doi:10.2196/jmir.5876 PMID:28062391

Simou, E., & Koutsogeorgou, E. (2014). Effects of the Economic Crisis on Health and Healthcare in Greece in the Literature from 2009 to 2013: A Systematic Review. *Health Policy (Amsterdam)*, *115*(2-3), 111–119. Retrieved May 13, 2017. doi:10.1016/j.healthpol.2014.02.002 PMID:24589039

Stroetmann, K. A., Artmann, J., & Stroetmann, V. (2011). Developing national eHealth infrastructures – results and lessons from Europe. *AMIA ... Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium*, *2011*, 1347–1354. Retrieved March 4, 2017 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3243126/ PMID:22195196

Tang, P. C., Ash, J. S., Bates, D. W., Overhage, J. M., & Sands, D. Z. (2006). Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption. *Journal of the American Medical Informatics Association*, *13*(2), 121–126. Retrieved May 16, 2017 from https://doi.org/10.1197/jamia.M2025 doi:10.1197/jamia.M2025 PMID:16357345

World Health Organization and International Telecommunication Union. (2012). *National eHealth Strategy Toolkit, Overview.* Retrieved March 4, 2017, from http://www.who.int/ehealth/publications/overview.pdf

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 7(1); edited by Anastasius Moumtzoglou; pages 16-30; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 73
Ethics of Electronic Health Record Systems

Brian J. Galli
Long Island University – Post, USA

ABSTRACT

This article describes how healthcare and IT are combatting the ethical implications of electronic health records (EHRs) in order to make them adopted by over 90% of small practices. There is a lack of trust in EHRs and uneasiness about what they will accomplish. Furthermore, security concerns have become more prevalent as a result of increased hacker activity. The objective of this article is to analyze these ethical issues in an effort to eliminate them as a hinderance to EHR implementation. As of now, 98% of all hospitals use EHRs. Between 2009 and 2015, the government allocated money and resources for incentive programs to get EHRs into every healthcare providers' office. During this time period, over $800 million dollars facilitated EHR implementation. Using this as a tool EHRs negative perception can be revitalized and combated with the meaningful use program. This article will highlight the ethical implications of EHRs and suggest ways in which to avoid them to make EHRs available in every healthcare provider.

INTRODUCTION

Information technology (IT) and systems (IS) are now very well integrated into aspects of today's society. The various technologies and systems that deal with an abundance of information yield great power, but also raise various issues. Unfortunately, many people exploit technology for personal gain, at the cost of others. This technology element makes ethics an extremely important aspect in IT.

Ethics in IT, also known as information ethics, concentrates on the creation, control, and use of information technology, and how it relates to the principles that govern society. The best practice to prevent technology misuse is to instantiate standards that avert such issues.

Creation is the development process that a technology employs. This can be the development of a new product or the implementation of an IS or IT. Project management is the most important aspect of creation. If the team doesn't utilize good tactics, they may overlook harmful issues. Control relates to the

DOI: 10.4018/978-1-7998-1204-3.ch073

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

software or technology distributor, maintenance of an IT or IS, and information management. Control focuses on ethical codes, legal or implied, that employees follow.

The final aspect is the use of IT or IS. Users cannot abuse privileges, especially if it is access to trade secrets or personnel files. Similar to control, use has a foundation in the company's code of ethics and corporate environment. Ethical breaches lead to legal, financial, and social difficulties. When a new IT or IS introduced, ethical implications arise concerning creation, control, and/or use; it is crucial to comprehend what ethical implications may arise.

The paper is organized as follows: the background section presents the needed foundational knowledge on the subject matter; it also presents the objective of the study. The research methods section presents how the methods used to identify and analyze the articles that served as the basis for achieving the objective of the study. The findings section outlines the analysis and findings from the execution of the research methods. The next section presents the discussion of the findings and analysis executed from the study; it also presents recommendations that should be used when applying ethics and EHRs in project/IT environments. The final section outlines the general conclusions and potential next steps as a means to close out this research study.

BACKGROUND

The healthcare industry must address all ethical implications in technology since it deals with highly sensitive information. As a result, a culture of ethical practices and standardized processes is nurtured. Electronic Health Record (EHR) systems are one technology that complies with ethical standards because of the way in which it handles patient information (Waegemann, 2003). Waegemann (2003) explains the purpose of EHR systems:

The concept of an EHR--electronic storage and instant availability of information to authorized practitioners--is often combined with the advantages of an electronic healthcare system, including enhanced access to medical information and greater efficiency. EHR promoters even claim that full access to health information might bring cures for certain diseases, such as AIDS. Healthcare is getting more complex every day. Today, multiple specialists are involved in most patients' healthcare, and paper records cannot keep practitioners completely informed. (p. 1)

EHRs play a pivotal role in keeping practitioners informed and connected. Its data can be collated and analyzed on a grander scale, thus enabling researchers to view and create trends about various health issues. It wasn't until recently that EHRs became more widespread and accepted. According to Tripathi (2012), EHRs have undergone vast transformation in the past 50 years. Predecessors to EHRs existed in the 1960s and early '70s. In the 1990s, the modern EHR came into fruition and practice. EHRs were slowly implemented because a lack of standardization in the market and little motivation from providers (Waegemann, 2003).

In 2009, EHRs experienced a substantial boom in innovation and adoption as a result of the Health Information Technology for Economic and Clinical Health (HITECH) Act (Tripathi, 2012). The HITECH Act contains a provision entitled the American Recovery and Reinvestment Act (ARRA), which add an incentive program and funding assistance to eligible facilities (CMS, 2016). Both HITECH and ARRA "motivated a fragmented customer base to act more like a single customer with coherent demand," thus

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

setting the stage for market standardization (Tripathi, 2012). Even with such changes, it is challenging for HER benefits to become more widespread. EHRs face ethical implications in healthcare and information technology, as discussed in the proceeding sections.

Study Objective

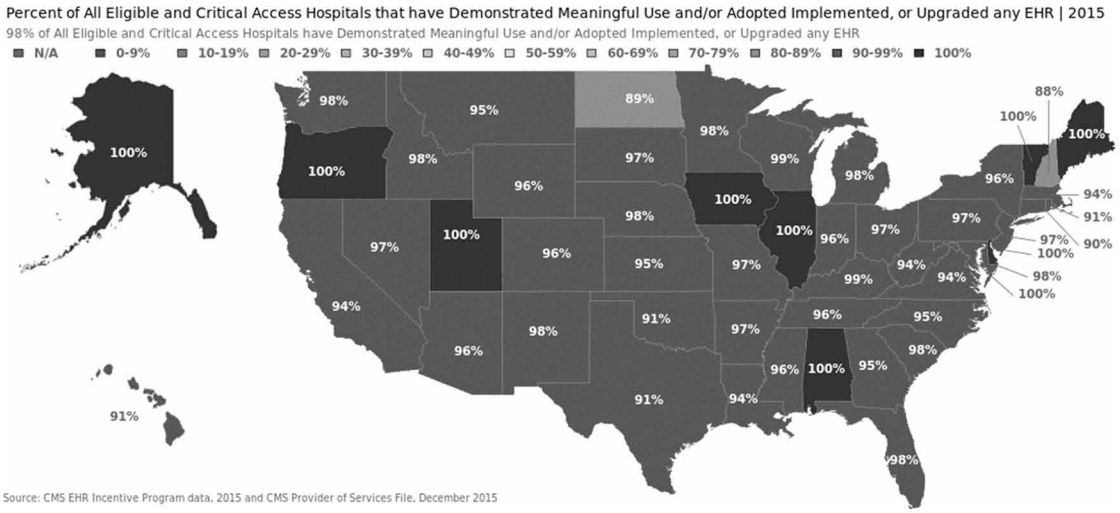
Healthcare faces important ethical questions because of the strict standards guiding the industry. Based on Figure 1, of all the hospitals that are eligible for assistance through HITECH, 98% demonstrated meaningful use of health IT and/or EHR implementation (Tripathi, 2012). Office-based health care professionals implementing EHRs experience more concern in maintaining and/or implementing EHRs.

As it is seen in Figure 2, only 57% of office-based health professionals use health IT and/or implement EHRs (Anderson, 2007). This affirms the notion that there is more concern in maintaining EHRs in office-based locations. Juxtaposed to the 98% of hospitals, there is an implication that either office-based healthcare professionals don't rely on the advanced patient information management an EHR provides, or there are roadblocks preventing them from implementing. Such roadblocks can translate to various ethical issues, as this paper suggests (Waegemann, 2003).

Office of the National Coordinator for Health Information Technology. (2016, February). 'Hospitals Participating in the CMS EHR Incentive Programs,' Health IT Quick-Stat #45. dashboard.healthit.gov/quickstats/pages/FIG-Hospitals-EHR-Incentive-Programs.php.)

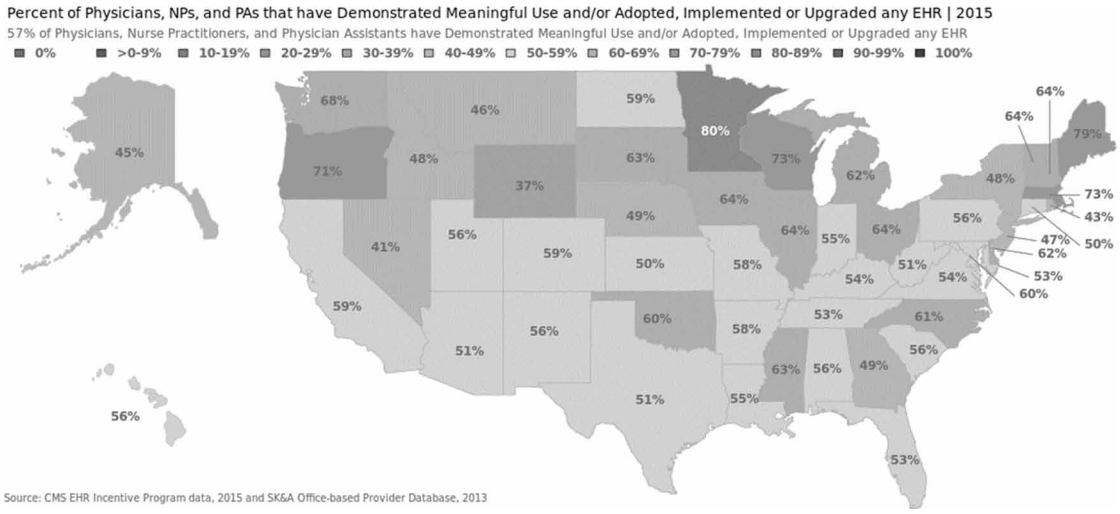
Some ethical issues can have a positive impact, such as finding and correcting care inequalities across racial or gender lines. Another positive implication is to more accurately diagnose illnesses via collective data. However, there are also negative ethical implications. This paper will review the clouded ethical nature of EHRs and Health IT by analyzing contested issues, positive and negative, surrounding EHRs and key ethical principles. Before advancing to the findings in this research, it is important to comprehend what research methods were implemented.

Figure 1. Hospitals Participating in the CMS EHR Incentive Programs (Source: EHR – Electronic Health Record, CMS – Centers for Medicare & Medicaid Services



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Office-based Health Care Professional Participation in the CMS EHR Incentive Programs (EHR – Electronic Health Record, CMS – Centers for Medicare & Medicaid Services, NP – Nurse Practitioner, PA – Physician's Assistant Office of the National Coordinator for Health Information Technology. (2016, February). 'Office-based Health Care Professionals Participating in the CMS EHR Incentive Programs,' Health IT Quick-Stat #44. dashboard.healthit.gov/quickstats/pages/FIG-Health-Care-Professionals-EHR-Incentive-Programs.php*.)*



RESEARCH METHODS

In order to perform a study on this topic, a systematic literature search (traditional narrative review) was performed. The traditional narrative review synthesizes the primary literature and explores potential heterogeneity. In this study, the review helped perform a literature review on the current fields of study and identify the potentially valuable pieces.

The traditional narrative review involved searching for literature in a 10-year window due to the main reason that EHR systems have become more prevalent in the last 10 years with the advancement in technology in the medical/healthcare field. A previous investigation into this topic also revealed that most studies into ethics of electronic health records have occurred in the previous 10 years, which added justification to looking back 10 years.

Our study's search strategy, following Stuck et al. (1999), consisted of three stages: first Scopus, ISI, Elsevier, and other universal research databases were used; second, we checked the references in existing articles; third, we asked experts for their recommendations. To run the queries within the online databases, keywords and phrases were explored (note: no filters were used and all combinations of keywords and phrases were considered). Our systematic search was conducted using selected general keywords (including: Ethics, Electronic Health Record, Healthcare, Information Technology, Information Systems).

Executing the searches returned 60 pieces of literatures (N = 60). We ended the search at saturation point - in other words when, combining the different keywords and adding new ones, we obtained articles already selected. To filter this large swath of literature, a meta-analysis was performed to filter out and identify the critical pieces. Furthermore, the meta-analysis was valuable in identifying potential gaps. Meta-analysis is a technique to aid going more in-depth into previous research. It is a systematic

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

method for combining relevant quantitative and qualitative data from different studies to develop one conclusion with greater statistical power. The meta-analysis conclusion is statistically stronger than the analysis of any single study because it incorporates an increased number of subjects, more diversity, and accumulated results.

From the original pieces of literature, the databases and searches were performed again using the authors' names. Using a search filter refined this analysis; the aim of a search filter is to unearth information of interest from a database. A simple filter for a query consists of single or combined words entered into the search parameters. Within search parameters, Boolean algebra is used to form relationships to further refine the output of query requests. The most commonly used Boolean functions are "AND", "OR" and "NOT."

For search filter development, two aspects need to be considered and balanced. They are sensitivity and precision (Taylor et al. 2003; Popay et al. 2004; Vaughan, 2004). Sensitivity is the ability of a filter to find all relevant material in the database, while precision is its ability to reject irrelevant material. A filter with a high emphasis on sensitivity will tend to include less and irrelevant material whereas a higher precision will more often reject some potentially relevant information. That is why it is crucial that both aspects be balanced to ensure that valuable information is not filtered out, but also so that irrelevant information is not filtered in.

The study investigator's approach was to develop a filter with a high level of sensitivity and then make further adjustments towards a desirable precision. Important material search listings obtained from early versions of the filter were examined. Then, common and key terms within this material (and in relevant cited literature) were incorporated to develop the filter's sensitivity. It was made sure that the filter incorporated a wildcarding of some terms.

From the precision perspective, examining common categories of irrelevant material returned by the more developed filter made the precision adjustment. After additional experimentation, the structure of the filter was adjusted based on the usage and placement of the "AND" operator and bracketing of the "OR" function. In practice, the combination of the specific requirements linked by the "AND" operator largely excluded irrelevant material to a workable level. For this study, the search filter included an "AND" relationship and "OR" (with brackets) to search for keywords and phrases. The filter developed specifically for this study was then applied to the search parameters of each database utilized in the initial search.

After literature was identified from the first set of queries, it was possible to examine the reference sections of each identified study. Therefore, more related studies are found. In some ways, the filter results demonstrated limited success in finding relevant materials in the wider gray literature. Gray literature is any material that is not commercially published. It includes working papers, business documents, government reports, educational reports, institutional reports, dissertations, technical reports, and conference proceedings. Gray material is found through manual searches, involving examination of reference lists from relevant studies obtained in searches, qualitative studies, relevant journals, books, and articles. This involved manual examination of relevant journal indexes.

Manual searching incorporates general and technical oriented Internet search engines and databases. Some include Google, British Library, US Library of Congress, Yahoo, and National Technical Information Service. As a final measure of robustness, to add in the systematic review of the available and relevant material, based on critical literature obtained, the study utilized direct correspondence with authors identified from the above processes. In so doing, additional information and studies could be noted to further support the objective of this paper.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Overall, the meta-analysis helped reduce N of 60 relevant studies to a more condensed and rich sub-sample size of 35 pieces (n=35), with 58.3% of the total literature found. From the papers identified, to further identify the themes and concepts that are presented in this paper, textual analysis was performed. This is a method used to describe and understand the traits of the 35 papers identified from the research literature review. A textual analysis enabled mapping and describing the content, functions, and structure of the literature identified.

Within the context of textual analysis, one of the common methods used is known as content analysis, in which a researcher used the literature to identify, process, enumerate, and analyzes specific themes and specific messages as well as message characteristics embedded in these texts. In this study, both qualitative content analysis (i.e. the meanings associated with messages as well as the number of times message factors occur) and quantitative content analysis (a systematic and step-by-step procedure that is used to address research questions/test hypotheses) were utilized.

Once qualitative and quantitative content analysis tools were complete, the results were categorized into specific themes and patterns through the use of affinity diagramming (a tool to organize ideas in a logical manner). The results of the affinity diagramming served as the basis for the themes and sections outlined and discussed throughout the paper. From the filtering and affinity diagraming process, the 35 articles were further reduced to 20 that studied the scope of the factors under scrutiny.

We limited our study selection to works that clearly explore the relationship or associations between the variables in the scope of our work. Amongst them, only the studies that provide sufficient information to assess a causal relationship between the factors being studied were included. Furthermore, we decided to exclude papers that did not explore the association between the factors being studied; the reason being that in these papers do not seek to understand any of the relationships or causal impacts as well as casual factors between the variables being explored in the study. These papers/studies would not measure/study the effect that our research study sought to explore.

At the end of this next-phased selection and filtering process, we came up with a reduced set of 20 papers on the variables being explored in the study. Some authors presented an analysis of more than one independent sample or studied more than one transition in the same paper. These estimates were treated as independent and were all included into our analysis. We accepted the estimates from final models only. When the same author published two papers using the same dataset and the same model specification, an average estimate was calculated based on the reported outcomes. However, when the same dataset was used, but a different model was estimated, we included both estimates in order to avoid the possibility of a study selection bias.

Overall, the selection of studies began with an N of 60 potential articles that could be used in this research. But the complete search procedure gave our study a total of 20 (33.33% of the of original population of selected papers) papers/studies that previously explored the variables being explored in the study. Thorough review of the 20 identified papers/articles showed that these pieces of literature varied in the focus and depth of the study of ethics in relation to healthcare and IT.

From a descriptive and trait perspective, the compilation of these papers showed that the relation of ethics in the area of healthcare and IT involved three major themes/traits, those being: organizational, legal, and ethical principles. The 20 studies/papers identified for this paper are categorized based on theme below in Table 1 (note: some studies overlapped into more than one category based on the research it pursued, but each study was categorized into 1 theme that best describes the study). Table 1 shows the categorization and frequency of each of the three themes/traits identified from the research execution.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Meta-Analysis Results by Identified Category

Category	Number of Articles (Frequency)	Article/Paper Authors
Organizational	8 (40% of total articles)	Ozair et al., 2015 Tripathi, 2012 Menachemi & Collum, 2011 Anderson, 2007 Taylor et al., 2003 Popay et al., 2004 Vaughan, 2004 Kazley et al., 2014
Legal	7 (35% of total articles)	Waegemann, 2003 CMS, 2016 Stuck et al., 1999 Palvia et al., 2015 Reynolds, 2015 Health IT, 2015 Cortelyou-Ward, Noblin, & Martin, 2011
Ethical Principles	5 (25% of total articles)	Ben-Assuli, 2015 HHS, 2015 Hodge et al., 1999 Layman, 2008 Noblin, Cortelyou-Ward & Ton, 2011

From a relational and causal perspective, review of the 20 identified papers/articles indicate that these pieces of literature analyzed ethics and healthcare and IT with different statistical methods (some not at all) as well as in depth of the study, especially in the scope of the amount of variables/factors explored in the study. The use of statistical analysis and exploration of different variables/factors in these studies helps to add value and depth to the conclusions drawn from our research study.

Table 2 presents a summary of the statistical methods the 20 studies used in their research. Table 3 presents a summary of the number variables/factors studied in each of the 20 studies, which shows the depth that the study explored.

The next section presents the findings of the research methods in the scope of these major traits and from the statistical methods used to perform the previous studies.

FINDINGS: ETHICS IN HEALTHCARE AND IT

Ethics is a set of moral actions that a group adheres to. It is applicable to healthcare and IT. To avoid confusion as to what the accepted right and wrong behavior is, codes, standards, or laws will define right and wrong for perceivable situations. In order to understand ethical issues concerning EHRs, ethics ought to be broken down further into the following characteristics: organizational, legal, and ethical principles.

Organizational

One of the largest bodies with the most influence over EHR implementation is the government. While other bodies come close to this influence, after 2009, the power shifted from the market to the government. The power shift was due in part to the Health Information Technology for Economic and Clinical

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Meta-Analysis Results by Statistical Analysis Method

Statistical Method	Number of Articles (Frequency)	Article/Paper Authors
Regression	6 (30% of total articles)	Ozair et al., 2015 Tripathi, 2012 Menachemi & Collum, 2011 Popay et al. 2004 Vaughan, 2004 Kazley et al., 2014
ANOVA	4 (20% of total articles)	Waegemann, 2003 CMS, 2016 Stuck et al., 1999 Cortelyou-Ward, Noblin, & Martin, 2011
Q-Test	3 (15% of total articles)	Ben-Assuli, 2015 HHS, 2015 Hodge et al., 1999
t-Test	5 (25% of total articles)	Anderson, 2007 Taylor et al. 2003 Palvia et al., 2015 Reynolds, 2015 Health IT, 2015
Chi-Square Test	2 (10% of total articles)	Layman, 2008 Noblin, Cortelyou-Ward & Ton, 2011

Table 3. Meta-Analysis Results by Number of Variables Studied

No. Factors Studied	Number of Articles (Frequency)	Article/Paper Authors
1	4 (20% of total articles)	Taylor et al., 2003 Popay et al., 2004 Vaughan, 2004 Kazley et al., 2014
2	4 (20% of total articles)	Waegemann, 2003 CMS, 2016 Palvia et al., 2015 Reynolds, 2015
3	4 (20% of total articles)	Ben-Assuli, 2015 HHS, 2015 Hodge et al., 1999 Health IT, 2015
4	4 (20% of total articles)	Ozair et al., 2015 Tripathi, 2012 Noblin, Cortelyou-Ward & Ton, 2011 Stuck et al., 1999
5	3 (15% of total articles)	Menachemi & Collum, 2011 Anderson, 2007 Layman, 2008
6	1 (5% of total articles)	Cortelyou-Ward, Noblin, & Martin, 2011

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

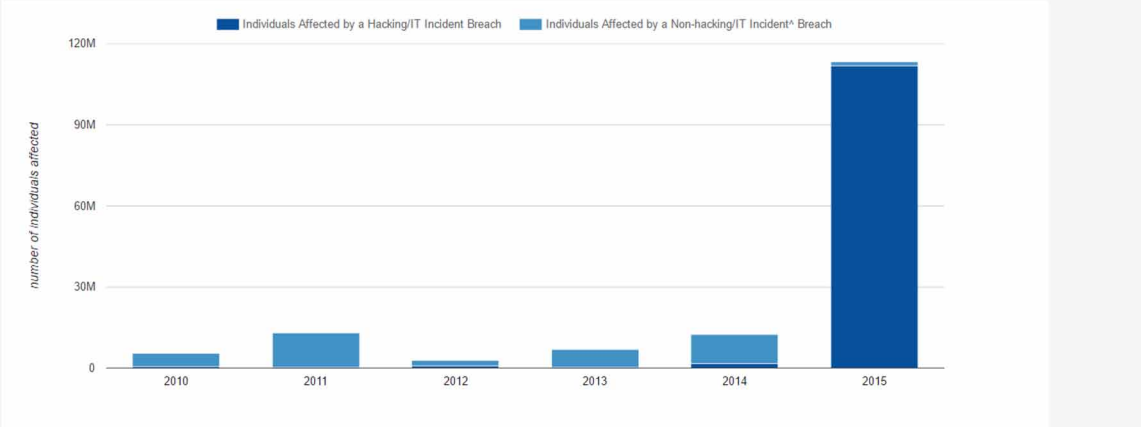
Health Act (Ben-Assuli, 2015). This was not an unethical act intended to disenfranchise involved parties. The government merely noticed that EHR would become a future necessity. However, there was a lack of motivation to make it a priority because, despite the benefits, implementation is time consuming and costly (Ozair et al., 2015).

The HITECH Act instantiated a program through the CMS that would incentivize healthcare providers to get on board with EHRs. This, in addition to getting 98% of hospitals on EHRs (according to Figure 1), helped move the market towards standardization (Tripathi, 2012). HITECH, and its child act ARRA, set requirements that healthcare institutions comply with. In order to qualify for and get compensation, the EHR system must meet certain criteria that are indicative of standardization, according to the acts (Menachemi & Collum, 2011).

Prior to HITECH, EHR developers implemented their own systems that were not linked with each other. Furthermore, there were no security standards to protect against millions of security breaches each year (see Figure 3). The EHR vendors, only after the government intervention in 2009, began working in a more concise manner to find solutions and build security that facilitated different systems to work together (Kazley et al., 2014). Prior, there was a "lack of ability to exchange clinical data with laboratories and hospitals," and this proved to be "a major barrier for smaller physician practices" because while hospitals have their own facilities, it is hard for small practices to get access to the facilities (Anderson, 2007). Anderson (2007) suggested that a solution to sharing barriers is to certify vendor EHR programs. Such a solution requires a form of standardization, which HITECH creates.

Another issue vendors caused was the "cost barrier," which dissuaded small practices from adopting EHR systems in the first place (Anderson, 2007). There is viable good solution for this issue, since vendors allocate resources to create a product and expect to be paid for it. It is hard to discern if the costs are overpriced, because vendors get away with such a grey area. HITECH and other initiatives have invested around $800 million toward EHR adoption since 2009, which has since then grown (Palvia, Jacks, Brown, 2015).

Figure 3. Breaches of Unsecured Protected Health Information. Number of individuals affected by protected health information breaches: 2010 – 2015



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In spite of the increased spending and governmental initiatives, many healthcare providers are still reluctant to adopt EHRs. This lack of urgency is attributed in part to the issues encountered at various stages of the implementation process in an organization (Palvia et al., 2015). Based upon Figure 2, while there have been significant increases in hospital adoption of EHR, this is not reflected in small clinics, because of implementation.

Many issues stemming from the implementation process include the stakeholder not feeling aligned with the benefits of the EHR. It is possible that stakeholders will back out and not adopts implementation (Palvia et al., 2015). The lack of urgency to adopt EHRs comes from the stakeholder's tendency to not be involved in the implementation process. They ought to be more involved. The lack of involvement can provoke a disconnect between the stakeholder and the EHR system. It is important to note that the stakeholder's perception of the systems use influences their desire to remain involved. Stakeholders' genuine interest is imperative for successful adoption.

Finally, something that is a potential point of contention between the vendor and the stakeholder incorporates security and privacy of the EHR system (Waegemann, 2003). Organizational characteristics in hospitals and smaller healthcare facilities have an impact on the adoption and implementation of EHRs. It raises similar ethical implications to the proceeding characteristics.

Legal

Ethics pertaining to legal issues are prevalent in healthcare on both the medical and vendor or IT side. Malpractice and negligence are unfortunate occurrences to be dealt with, which decreases the interest of facilities to pursue EHR implementation. When an EHR is installed, negligence can occur in two ways. The IT professional may improperly implement a part of the system, causing extended downtime or file corruption. If this is not a breach of duty to care, then it is not legal negligence (Reynolds, 2015). However, mistakes such as this one maintain a measurable impact, especially on the project management team or process. If the offending party is brought to court and the act is ruled a breach of duty to care, and then there are potential legal ramifications for the individual and company (Anderson, 2007).

The second issue in which negligence may occur is after an EHR is installed. If it is not properly maintained and updated, maintenance issues may surmount (Menachemi & Collum, 2011). The risk of losing patient files or having a security breach is more prevalent because of this form of negligence. This is one of the barriers and reasons that many practitioners in the healthcare industry may not trust EHR systems.

Another issue that EHR systems face is reliability, which is especially true in facilities with heavy involvement in emergency situations (Menachemi & Collum, 2011). EHRs need to be quality, safety-critical systems since they handle sensitive information that relates to life or death situations. Neglecting to design and implement safety-critical systems is not permissible. Instead, the facilities ought to use special development methodology to ensure the system is within acceptable risk factors and that vendors are protected for liability (Reynolds, 2015).

EHRs produce various implications for medical professionals. The EHR documents information without prejudice. Therefore, if a medical staff member makes a mistake, liability of the staff member and the entire facility are increased (Ben-Assuli, 2015). It is possible, however, that with EHRs, staff member will make less mistakes, thus reducing the number of malpractice claims. It is deemed that "this reduction is potentially the result of increased communication among caregivers, increased legibility and completeness of patient records, and increased adherence to clinical guidelines" (Menachemi

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

& Collum, 2011). Medical staffs are held accountable for their actions, regardless of their facility or resources. With EHRs, their accountability remains, but there is a level of forgiveness for human error, because the system can catch innocent mistakes before causing issues.

In the healthcare industry, privacy is a crucial domain. Patient-physician confidentiality is one of the most prominent aspects of healthcare privacy, in which a doctor will not, and cannot, reveal patient information. EHRs must ensure that patient privacy is maintained. EHRs are generally quite secure, and personal medical information is only accessible by authorized practitioners (Kazley et al., 2014).

The Health Insurance Portability and Accountability Act (HIPAA) of 1996 promote the compliance of EHRs with patient confidentiality. HIPAA's coverage of an individual's rights over the privacy and security of their health information consists of two rules, The Privacy Rule and The Security Rule. The two rules are outlined as follows (HHS, 2015).

The Privacy Rule, a Federal law, gives you rights over your health information and sets rules and limits on who can look at and receive your health information. The Privacy Rule applies to all forms of individuals' protected health information, whether electronic, written, or oral. The Security Rule is a Federal law that requires security for health information in electronic form.

HIPAA contains a rule that adds provisions to the HITECH Act to strengthen privacy and security measures (HHS, 2015). These provisions must be met as part of the EHR incentive program in order to protect patient information confidentiality. Some are concerned about the legitimacy of EHR privacy since many EHRs utilize web interfaces (Anderson, 2007). To someone who does not fully comprehend the technology, this is a valid concern. However, EHRs that incorporate a web interface is often more secure because of this identified issue, which would have a greater likelihood of appearing.

While privacy and confidentiality are fairly easy to maintain, use of an EHR "increases opportunities for invasions of privacy by dispersing health data through interceptable telecommunications, thus magnifying the risks of breaching patient-physician confidentiality." (Hodge et al., 1999). Such intrusions might appear contradictory to the legal and ethical safe guards that have been implemented through acts. The contradiction arises, however, because there is a thin line between information security and privacy. There is no doubt that once a security breach occurs and patient information is compromised, privacy will be invaded. Unfortunately, once private information is connected to the World Wide Web, those individuals become easier targets for different forms of malicious attacks.

In 2015, the security breach affecting the most people was "Hacking/IT Incident". In previous years, security breaches were less effective and precipitated from non-hacking, unintentional scenarios (Figure 3). During the same time period, the most common attack was "Theft", followed by "Unauthorized Access/Disclosure." Table 4 demonstrates these attacks.

It is evident that as technology advances, so do the risks involved with using technology. There are several regulations in place to attempt bolstering security as a means to protect patient privacy and confidentiality. However, with the exception to the statistical outlier in 2015, most security breaches in healthcare occur because of the lack of training in basic cyber security that healthcare employees are given.

Many security threats also precipitate from stolen devices or improperly managed access. Hodge et al. (1999) described the impediment of health privacy, security, and technology by suggesting changes to HIPAA or other health privacy legislation. Hodge et al. state, "Health-related computer technology requires a weighing of competing interests toward devising effective policies to protect the public." The legal characteristics of implementing EHRs concerning negligence, liability, privacy, and security are reasons that EHR implementation has been slow in America (Anderson, 2007). The proceeding section will further discuss the ethical principles involved in implementing EHRs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Count of affected individuals by the type and source of information breach

	2010	2011	2012	2013	2014	2015
Type of Information Breach						
Hacking/IT incident	568,358	297,269	900,684	236,897	1,786,630	111,812,172
Improper disposal	34,587	63,948	21,329	526,538	93,612	82,421
Loss	924,909	6,019,578	95,815	142,411	243,376	47,214
Theft	3,691,460	4,720,129	927,909	5,397,989	7,058,678	740,598
Unauthorized access/ disclosure	130,106	118,444	338,767	383,759	3,019,284	572,919
Other breach	158,593	13,981	503,900	254,305	413,878	--
Source of Information Breach						
Desktop computer	246,643	2,042,186	81,385	4,348,129	2,378,304	316,226
Electronic medical record	803,600	1,720,064	136,751	40,196	121,845	3,948,985
E-mail	8,050	3,111	294,308	58,847	519,625	583,977
Laptop	1,507,914	405,873	575,529	1,023,181	1,273,612	391,830
Network server	665,123	613,963	921,335	320,127	7,253,441	107,252,466
Paper/Film	204,966	103,711	198,409	575,076	590,352	229,743
Portable Electronic Device	29,714	1,516	124,978	154,877	141,110	209,558
Other source	2,058,166	8,259,368	455,709	422,381	343,537	322,539

Office of the National Coordinator for Health Information Technology. (2016, February). 'Breaches of Unsecured Protected Health Information,' Health IT Quick-Stat #53. dashboard.healthit.gov/quickstats/pages/breaches-protected-health-information.php.

Ethical Principles

The final characteristic of ethics in EHR implementation is general ethical principles, which includes social responsibility. Healthcare and IT have a responsibility to ensure that its code of ethics is upheld, despite rapid industry changes. In order to understand this ethical identity crisis, ethical principles that define EHRs ought to be explained. Layman (2008) chose two ethical principles, beneficence and autonomy, to outline EHRs intentions and failures.

Beneficence relates to the intentions of EHRs. EHRs' intention is to improve the quality of healthcare. This is considered a good and valuable intention. Government programs and laws, such as HITECH, focus heavily on combating issues that EHRs experience before regulation or standardization existed. EHRs exist to help others, and therefore cannot be considered malicious by nature. Unfortunately, it is important to note that however good EHR intentions may be, "beneficence is not assured" (Layman, 2008).

When EHRs face heavy criticism surrounding security, many will not call it a good system. As mentioned in the previous section, security breaches often result from stolen devices that do not have proper security protection installed. Furthermore, such theft, outside of 2015, is one of the highest causes for security breaches (Ben-Assuli, 2015).

Another reason many will not consider EHRs as a good system is because its users are poorly trained. This directly endangers the system's security and privacy. The patients' confidentiality is put into the hands of staff that do not know how to manage and protect the information. As a result, the true good

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

that EHRs can bring to patients and healthcare are ignored. As one knows, successful implementation can help cure diseases and save many lives (Ben-Assuli, 2015). There is a visible reduction in costs when advanced EHR systems are implemented, as demonstrated in Figure 2. Furthermore, EHRs can reduce malpractice. Even though they are supported by government initiatives, they are still questioned.

The next ethical principle, per Layman (2008), is autonomy. EHRs are known for being potentially unethical because of who gains the right to control patients' health information. Even with HIPAA regulations, some feel uncomfortable about which health data is released to. This uncomfortable feeling is where the concept of informed consent becomes crucial. Informed consent closely follows the idea that "information of a patient should be released to others only with the patient's permission or allowed by law" (Ozair et al., 2015). Therefore, any secondary use of a patient's health information that Palvia et al. (2015) discusses would be deemed illegal. The concept of illegality contradicts the good and valuable aspects behind data sharing. If people generally feel uncertain about sharing health data, then EHRs may not be considered good in nature anymore. However, a study found that "50% of respondents were concerned that their records would be linked outside their primary doctor's office, although 79% of them recorded that they believed making health records electronic was a good idea" (Layman, 2008). Many people believe that it is ethically incorrect to transfer information without informing a patient.

Autonomy relies closely on the concept of access. There is an added possibility that a minority group is denied access to certain care. The reason can be overtly racist or based primarily off economic status. EHRs offer the potential to view in-depth information that carries social, as well as medical, information about patients. However, this concept dances on a fine line. It can be seen as counter intuitive to the concept of autonomy, and thus perpetuating inequalities, especially if there is "public disclosure of personal health information of certain disadvantaged groups." (Layman, 2008).

Examining the issues inflicting EHRs through such ethical principles yields a different perspective on the qualities of an EHR. EHRs contain both positive and negative attributes. Healthcare and IT are ethically motivated fields. There is a professional and social responsibility to ensure that EHRs are not unethically abused (Palvia et al., 2015). The ethical characteristics outlined in the previous subsections demonstrate why it is challenging to implement EHRs. However, it is concluded that the positive ethical impacts of EHRs make it worth the implementation process.

In conclusion, this study found that the impact and relationship that ethics has with IT and IS in the context of healthcare revolve around three main traits. They include organizational, legal, and ethical principles. Therefore, there are some ethical implications and recommendations that a practitioner ought to be aware of as they execute the development and implementation of EHR systems. These are discussed in the following section.

DISCUSSION

Implication of the Ethics in Project Management and IT

EHRs are complex systems that require a great deal of planning and project management. Understanding the ethical issues can sway a decision in project implementation that directly impacts the EHRs and their effectiveness. When initiating the process to create an EHR system, communicating with the staff

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and patients is crucial to discern what each party expects from the system (Noblin, Cortelyou-Ward & Ton, 2011). It is important to take notice of concerns that can be addressed to reduce the possibility of future ethical issues from arising.

Once the vision and issues are identified, selecting a vendor is the next step. EHRs need to function as reliable, high quality systems. This comes as a job of the vendor. Privacy and Security are two ethical issues that are key in the implementation of an EHR. It is important to know whether or not a vendor is reliable and has good security credibility. With EHRs, cost saving is not the goal, as usability, reliability, and desire functionality are primary components paralleled in importance with ethical issues (Kazley et al., 2014). If patients are concerned about what will happen to their information once it is stored with the vendor or data management party, it is a good idea to negotiate for an informed consent policy that permits patients to opt out of data aggregation.

As Table 5 and Table 6 demonstrates, there is evidence to support the notion that a good EHR system will save money compared to a lesser one. It is a necessity that during implementation, "constant tracking of project progress and tight regulation of project tasks, personnel, communications, and budget" are maintained (Noblin et al., 2011). Even when money should not be a crucial factor, it is in this case, because controlling the budget along with strict monitoring and quality control is part of successful EHR implementation. It is also important to maintain quality control of documentation and data entry.

As part of healthcare implementation, security over who has access to patient data is highly important. This will help avoid ethical issues from occurring in the future and uphold the privacy and ethical laws put in place. EHR implementation is a rigorous process, but the issues that may occur should implementation be achieved poorly are more drastic (Tripathi, 2012).

The ethical issues with security and privacy in EHRs translate to general information technology. Most of the issues that EHRs face occur from improperly trained, or even untrained employees, who are carless with their access and handling of sensitive data. EHR impediments set a good example in what to look for in IT security. It is usually difficult for a hacker to get past network security, but he/she can use other means to get credentials out of unsuspected employees, especially when they are poorly trained on using the new system. Thus, EHRs will provide a lesson in maintaining and ensuring privacy (Kazley et al., 2014).

Table 5. Cost Estimate Sensitivity Analysis

	Base Estimate Propensity Strata N = 5,163,794	Base Estimate Propensity Matched Sample N = 152,008
EHR cost/administration	$7207	$7103
Non-EHR cost/administration	$7938	$8005
Cost difference	$731[a]	$902[a]

P values were calculated using multivariable Log-linked Gamma Distributed Generalized Linear Model.
[a]*P* <.0001. *P* values were calculated using multivariable Log-linked Gamma Distributed Generalized Linear Model.
EHR indicates electronic health record.
Covariates include: patient age, gender, race, All Patient Refined Diagnosis Related Groups (APDRG) mortality and severity, neonatal or maternal status, private insurance coverage, Medicare or Medicaid coverage, Diagnosis Related Group (DRG) case mix group, and whether the patient arrived as a transfer, hospital teaching status, hospital urban location, hospital bed size, and hospital geographical region.

Source: Association of Electronic Health Records with Cost Savings in a National Sample. (2014)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Cost Estimate Sensitivity Analysis (Descriptive Data Proof)

	Total (N = 5,163,794) Mean (SD)	Advanced EHRs (N = 1,509,610) Mean (SD)	No Advanced EHRs (N = 3,654,184) Mean (SD)
Age, in years	48.3 (27.9)	45.9 (28.0)	49.3 (27.7)
Risk mortality	1.6 (0.9)	1.6 (0.9)	1.6 (0.9)
Risk severity	2.0 (0.9)	2.0 (0.9)	2.0 (0.9)
Cost per admission	$10,790 (17,857)	$10,203 (17,572)	$11,010 (17,993)
	N (%)	N (%)	N (%)
Medicaid	1,078,607 (20.9)	332,439 (22.0)	746,168 (20.4)
Medicare	1,899,951 (36.8)	4,979,650 (33.0)	1402301 (38.4)
Private insurance	1,717,720 (33.3)	529,941 (35.1)	1,187,779 (32.5)
Neonatal/maternal admit	1,177,819 (22.8)	364,631 (24.2)	813,188 (22.3)
Transfer into hospital	284,239 (5.5)	106,651 (7.1)	177,588 (4.9)
White	2,952,903 (57.2)	787,539 (52.2)	2,165,364 (59.3)
Black	587,325 (11.4)	206,575 (13.7)	380,750 (10.4)
Hispanic	623,272 (12.1)	171,195 (11.3)	452,077 (12.4)
Teaching hospital	2,525,001 (48.9)	970,582 (64.3)	1,554,419 (42.5)
Urban hospital	4,652,205 (90.1)	1,457,757 (96.6)	3,194,448 (87.4)
Small hospital	585,247 (11.3)	114,943 (7.6)	470,304 (12.9)
Medium hospital	1,239,321 (24.0)	386,845 (25.6)	852,476 (23.3)
Large hospital	3,339,226 (64.7)	1,007,822 (66.8)	2,331,404 (63.8)
Northeastern United States	1,405,749 (27.2)	573,844 (38.0)	831,905 (22.8)
Midwestern United States	668,842 (13.0)	303,385 (20.1)	365,457 (10.0)
Western United States	1,477,224 (28.6)	331,228 (21.9)	1,145,996 (31.4)
Southern United States	1,611,979 (31.2)	301,153 (20.0)	1,310,826 (35.8)

EHR indicates electronic health record.

From one end, IT ought to ensure that ethical practices involving personal information are upheld. This is important because once a culture is developed in which breaking policies is the norm; it is more challenging to change the culture. One of the final aspects of the privacy issue is that information may never be truly safe. Practitioners must know what is accessible, and how to keep tabs on potential issues that may arise at any time (Tripathi, 2012).

As one can see, a healthcare facility must comply with ethical practices in project management and EHR implementation. Compliance will ensure that in the future, ethical organizational, legal, privacy, and security implications do not occur. Pursuing the suggestions made in this section will provoke smaller and/or on-the-fence facilities to implement EHRs and reap its positive, ethical benefits.

Recommendations

As a project manager interested in implementing an EHR system, early planning and initiation within the bounds of what is known about available resources must be undergone. A recommendation "can be made to make the initiation and planning stages a solid foundation for the rest of the project," and it is hard to build on a careless base if awareness of requirements is lacking (Noblin et al., 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In the planning stage, the project team must create a clear vision of the purpose of the EHR system. Everyone must agree on this vision. All those involved must share it because that will ensure that the process can run smoother. If the involved parties do not share the same line of vision, then the team risks conflict and unsuccessful, unethical implementation of the EHR system.

A final recommendation is to assess who the stakeholders are. By involving the stakeholders, the team can understand their vision and opinions about the project. It may not align exactly with each other, but there will be valuable insight into how to get the support needed to move forward on a project (Menachemi & Collum, 2011). Furthermore, involving the stakeholders will ensure that the produced system satisfies all their needs with quality.

In some respects, ethics in EHRs and healthcare IT is considered a messy matter. It is important to look at the situation without bias in order to collect all views. This collection can be achieved during the planning/implementation phases. If security is a concern, a plan of action should be created to handle it immediately (Waegemann, 2003). Security should be one of the topics of interest from the beginning of the project. It will not be beneficial if the team waits for something to happen before trying to prevent the ethical implications discussed in this study from arising.

Proactive ethics training is one of the best ways to combat issues with regards to ethics in the organization, legal, privacy, and security characteristics. It is important to ensure that all people involved are cognizant of potential legal issues, whether concrete as HIPAA or more ambiguous. Ethics education, along with basic workplace and out of workplace security education, can go a long way to prevent ethical issues and potentially illegal ones (Waegemann, 2003). An education on ethics will also further promote smaller facilities to embark on EHR implementation.

CONCLUSION

EHR systems are fraught with potential ethical issues. Some are combatable, while others are less so. It is clear that ethical issues with regards to EHRs are not going to be diminished. EHRs are in a crisis right now, especially with regards to their level of security, so is it crucial that we attempt tackling the ethical issues before they appear, instead of waiting until the last minute.

EHR systems will eventually yield positive results once security and ethical implications are eradicated. The real roadblock to success and defeating the ethical issues that are apparent is thinking that there is no potential to change how the system functions. There is much support, especially through government initiatives, for EHR systems. Some of the support comes from a team being motivated in the right direction. It may also include involving stakeholders and restoring trust into the system. Solutions for issues surrounding autonomy and informed consent are already possible—they just require implementation. When EHRs are given the chance to develop and be implemented, their impact on the healthcare industry will be astounding.

REFERENCES

Aggarwal, P., Jamshed, N., Ozair, F. F., & Sharma, A. (2015). Ethical issues in electronic health records: A general overview. *Perspectives in Clinical Research*, *6*(2), 73. doi:10.4103/2229-3485.153997 PMID:25878950

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Anderson, J. G. (2007). Social, ethical and legal barriers to e-health. *International Journal of Medical Informatics*, *76*(5), 480–483. PMID:17064955

Arai, L., Petticrew, M., Popay, J., Roberts, H., Rodgers, M., Roen, K., & Sowden, A. (2004). *Guidance on the conduct of narrative synthesis in systematic reviews (Draft report from ESRC Methods Programme)*. Institute for Health Research, University of Lancaster.

Beck, J. C., Büla, C. J., Hohmann, C., Nikolaus, T., Stuck, A. E., & Walthert, J. M. (1999). Risk factors for functional status decline in community-living elderly people: A systematic literature review. *Social Science & Medicine*, *48*(3), 445–469. PMID:10075171

Ben-Assuli, O. (2015). Electronic health records, adoption, quality of care, legal and privacy issues and their implementation in emergency departments. *Health Policy (Amsterdam)*, *119*(3), 287–297. doi:10.1016/j.healthpol.2014.11.014 PMID:25483873

Brown, W., Jacks, T., & Palvia, P. (2015). Critical issues in EHR implementation: Provider and vendor perspectives. *Communications of the Association for Information Systems*, *36*(36), 45–68.

Centers for Medicare & Medicaid Services (CMS). (2016). Electronic health records (EHR) incentive programs. Retrieved from https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html

Collum, T. H., & Menachemi, N. (2011). Benefits and drawbacks of electronic health record systems. *Risk Management and Healthcare Policy*, *30*(4), 47–55. PMID:22312227

Cortelyou-Ward, K., Martin, J., & Noblin, A. (2011). Electronic health record project initiation and early planning in a community health center. *The Health Care Manager*, *30*(2), 118–124. doi:10.1097/HCM.0b013e318216eeff PMID:21537132

Cortelyou-Ward, K., Noblin, A. M., & Ton, S. (2011). Electronic health record implementations: Applying the principles of monitoring and controlling to achieve success. *The Health Care Manager*, *30*(1), 45–50. doi:10.1097/HCM.0b013e3182078b4f PMID:21248548

Dempster, M., Donnelly, M., & Taylor, B. J. (2003). Hidden gems: Systematically searching electronic databases for publications for social work and social care. *British Journal of Social Work*, *33*(4), 423–439. doi:10.1093/bjsw/33.4.423

Gostin, L. O., Hodge, J. G. Jr, & Jacobson, P. D. (1999). Legal issues concerning electronic health information: Privacy, quality, and liability. *Journal of the American Medical Association*, *282*(15), 1466–1471. doi:10.1001/jama.282.15.1466 PMID:10535438

HealthIT. (2016). Quick Stats. Retrieved from http://dashboard.healthit.gov/quickstats/quickstats.php

HHS. (2015). Health information privacy. Retrieved from http://www.hhs.gov/hipaa/

Kazley, A. S., Simpson, A. N., Simpson, K. N., & Teufel, R. (2014). Association of electronic health records with cost savings in a national sample. *The American Journal of Managed Care*, *20*(6), 183–190. PMID:25180501

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Layman, E. J. (2008). Ethical issues and the electronic health record. *The Health Care Manager*, *27*(2), 165–176. doi:10.1097/01.HCM.0000285044.19666.a8 PMID:18475119

Reynolds, G. W. (2015). *Ethics in information technology* (5th ed.). Boston, MA: Cengage Learning.

Tripathi, M. (2012). EHR evolution: Policy and legislation forces changing the EHR. *Journal of American Health Information Management Association*, *83*(10), 24–29. PMID:23061349

Vaughan, L. (2004). New measurements for search engine evaluation proposed and tested. *Information Processing & Management*, *40*(4), 677–691. doi:10.1016/S0306-4573(03)00043-8

Waegemann, C. P. (2003). EHR vs. CPR vs. EMR. *Healthcare Informatics*, *43*(2), 430–451.

This research was previously published in the International Journal of Information Systems and Social Change (IJISSC), 9(3); edited by John Wang and Jungwoo Lee; pages 53-69; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 74
Hierarchy Similarity Analyser–An Approach to Securely Share Electronic Health Records

Shalini Bhartiya
IITM, India

Deepti Mehrotra
Amity University, India

Anup Girdhar
Sedulity Groups, India

ABSTRACT

Health professionals need an access to various dimensions of Electronic Health Records (EHR). Depending on technical constraints, each organization defines its own access control schema exhibiting heterogeneity in organizational rules and policies. Achieving interoperability between such schemas often result in contradictory rules thereby exposing data to undue disclosures. Permitting interoperable sharing of EHRs and simultaneously restricting unauthorized access is the major objective of this paper. An Extensible Access Control Markup Language (XACML)-based framework, Hierarchy Similarity Analyser (HSA), is proposed which fine-grains access control policies of disparate healthcare organizations to achieve interoperable and secured sharing of EHR under set authorizations. The proposed framework is implemented and verified using automated Access Control Policy Testing (ACPT) tool developed by NIST. Experimental results identify the users receive secured and restricted access as per their authorizations and role hierarchy in the organization.

DOI: 10.4018/978-1-7998-1204-3.ch074

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

The inclusion of electronic media in healthcare domain resulted in transformation of paper-based health records into Electronic Health Records (EHRs). These records are accessed from medical to non-medical users at the same time across multiple hospitals and other related organizations. These records possess a high level of sensitivity where the patient demands utmost care allowing only relevant and authorized access of the records. Rules and policies are defined with respect to the roles and responsibilities of health providers. It may differ from hospital to hospital. With the heterogeneity involved in the organizational structure and access policies, sharing EHR increases the chance of confidentiality leakages and security breaches. Organizations, hospitals in this context, experience various challenges and security issues (Bhartiya & Mehrotra, 2014; Bai et al., 2014) in making the data available to the intended users.

In spite of similar roles of the users, the two hospitals may exhibit dissimilarity in their organizational hierarchies. Hence, to enable cross- organizational sharing of data, well-defined standards or mechanisms should be developed. The objective is to logically identify similarities between user's roles and authorities and make two disparate organizations interoperable. For ex. comparing two doctors, one in government hospital and the other in private hospital, their work distribution may not be the same and hence, the set of access rules must differ in both cases. A possibility exists where an access to the resource should be permitted to one and denied to other. The case reflects a huge challenge in determining and setting their access rights and privileges for secured sharing of sensitive EHR in-spite of wide gap in their services.

Basically, sharing of EHR is dependent on how best the two systems can achieve interoperability. HL7 has been addressing interoperability problems in sharing or exchanging EHRs for over two decades now. It (HL7 Standards, 2015; HL7 Evolution, 2010) provides a framework for exchange and sharing of EHRs for seamless integration between disparate healthcare organizations. The Standard is broadly divided into two categories – Version 2 (V2) and Version 3 (V3). Fast Healthcare Interoperability Resources Specification (FHIR), Release 1 (HL7 FHIR, 2015) combines the best features of HL7's V2, V3 and CDA product lines. It is a standard for exchanging healthcare data electronically. It simplifies implementation without compromising information integrity. Integrity is complemented with confidentiality and availability of data.

Interoperable sharing of EHRs exhibits another challenge -maintaining the confidentiality and privacy of sensitive health information without compromising its availability to the health providers. NIST (NIST 800-53, 2013) in collaboration with the Office of the Director of National Intelligence (ODNI), the Department of Defense (DoD), and the Committee on National Security Systems (CNSS) is establishing a unified information security framework for the federal government. This framework is said to provide a strong base for mutual acceptance of authorization decisions and facilitate information sharing between disparate organizations.

An access control framework allocates the access rights and privileges on the resources. A typical deployment of an access control system is a combination of three logical components: an access control model (ACM), policies, and enforcement mechanisms. A logical access control system requires validating individual's roles and responsibilities for assigning the access rights on the data. As each system is autonomous specifying access control policies according to the user and resource hierarchy of the organization, collaborating such organizations often result in policy conflicts (Huang and Liu, 2009), matching of undefined attributes or rules, thus, imposing a threat to relevant and meaningful disclosure of the information. Support from industry standard access methods would enable the disparate EHR-systems to collaborate and expand access to the data for providing quality (de la Torre-Diez et al., 2012)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

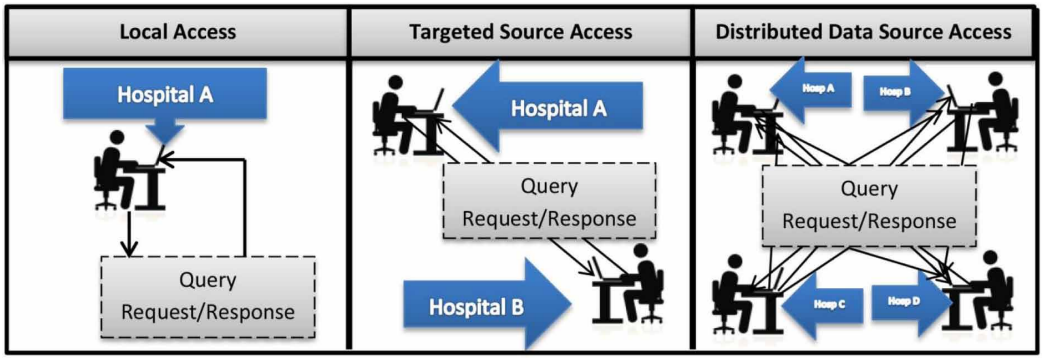
and timely care to the patient. The Standards and Interoperability (S&I) Data Access Framework (DAF) Initiative outlines the standards and profiles necessary for clinicians, providers and healthcare professionals to gain access to patient data within an organization and across organizations. DAF addresses three basic challenges-Who is accessing data, what and how the data is accessed and why the data is accessed?

Figure 1 identifies the different environments of information access requirements with respect to interoperable healthcare domain. DAF proposes to implement interoperable data access through a standard-based initiative delivered in two phases. Phase 1 would cater to the standard for local access via intra-organization query whereas phase 2 will focus on targeted access via inter-organization or distributed query. The standards selection is made keeping the modularity and substitutability into view. In order to enable the systems to interoperate, DAF specifies that organizations should implement consent requirements per their state, local, and institutional policies. However, there are no mandatory requirements for consent in DAF context. Moreover, access should be complemented with authorization control to provide a secure and reliable sharing (Parker & Wisley, 2009) of patient's health information with minimum or no change in the existing access control models. Hence, the aim of this study is to design an access control framework providing legitimate and only relevant access to EHR by deducing similarities between the roles of the users or health providers of disparate organizations.

Designing of ACM should strongly consider a comprehensive standards-based approach that enables current day interoperability capabilities and are cost efficient. Well-defined and standardized models should be encouraged, especially in cross-organizational environments. Due consideration should be given to functionality, interfaces, infrastructure, and IT support keeping in view the nature of users operating the system. A filtered mechanism must be employed for procuring the authorized access regardless of categorization of organizations, i.e. public, private or individual ownerships.

In this paper an ABAC based framework is proposed that integrates XACML policies by ranking each attribute on the basis of hierarchical distances between the user hierarchies of the healthcare organizations. The Hierarchy Similarity Analyser (HSA) algorithm generates security level (SL) that is assigned to users as per their positions in the organization hierarchy. SL is associated with the subject attribute on rules of access control policies. The policies are hence, fine-grained by associating authorizations based on Mandatory Access Control (MAC) rule. This allows a robust security control in limiting the amount of data to be accessible between two disparate health organizations. Lin et al. (2007) proposed a similarity-based policy filtering algorithm that identifies similar policies by calculating hierarchical distances between two or more attributes defined in the rule sets. Our framework uses a part of this al-

Figure 1. Data access between intra, inter and distributed hospitals



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

gorithm to calculate the similarity scores between two disparate user hierarchies. The similarity scores thus obtained are inputted to Hierarchy Similarity Analyser (HSA) that generates security level (SL) for each user in the given hierarchies. The SL obtains a clear relationship between the users with similar roles and similar or dissimilar responsibilities.

Access Control Policy Testing (ACPT) tool (Hwang et al., 2010) developed by NIST is used to verify the proposed framework. A comparative study by Li et al. (2015) reveals the significance and robustness of ACPT over other similar verification tools. ACPT bridges the gap between disparate policies by generating enforceable policies directly from policy models. The tool allows static as well as dynamic verification of the generated policies thereby helping in the identification and reduction of faults in the policies. Static verification is done through SMV (Symbolic Model Verification) model checker that checks given policy models to satisfy the user-specified properties. Dynamic verification is done by automatically generating and executing viable oracles through counterexample that illustrate semantic differences between the two policies. More specifically, each counterexample represents a request that evaluates to a different response when applied to the two policy versions.

This paper is divided into 5 sections. Section 1 details the problems in linking of heterogeneous health organizations with respect to access control policies that often contradict with each other arising confidentiality threats to sharable EHRs. Section 2 includes a literature review of related access control frameworks in other domains. Section 3 proposes the access control framework build on XACML structure. Section 4 deals with verification of the proposed framework using an automated tool ACPT. It takes up a case study to illustrate the working and testifying the framework. Section 5 concludes the results obtained justifying the viability of the proposed framework.

2. RELATED WORK

Various access models (Amini and Jalili, 2010; Hu et al., 2013; Thomas, 2007; Park & Lee, 2007; Karp et al., 2010; Hu & Kuhn, 2011; Miyagawa et al., 2014) have been proposed that promise to provide a well-integrated interoperable environment in their respective domain of study. Each of the model exhibit different architecture and features with respect to the access control policies.

The model proposed by Imine et al. (2009) deals with the latencies and resolves the policy conflict changes dynamically but allows any change to existing data-sets only at the administrator level. The users play a vital role in providing health services to the patient. Each user having any position in the organizational structure is equally crucial in the service-chain system of the organization. The organization experience frequent role-reversal of its users as per the demand and schedule. A domain-based approach (Wickramasuriya, 2004) to access control hides the sensitive data thereby acting as a filter at the object level. It establishes a security level on the domain of actors dynamically created and modified as per the need. Some organizations require reviewing the capabilities associated with subjects and their attributes defined through the access control policies. Access Control Lists work on the attribute of "identity" whereas RBAC works on the attribute of "role". The policies in ABAC expresses a complex Boolean rule set capable of evaluating many different attributes. The RBAC-based model (Power et al., 2011) captures the disparities of access control policies relying on specific constraint-checks to administer interoperability. With the advent of mobile-based (Rodrigues et al., 2013a) and cloud-based solutions (Fernández-Cardeñosa et al., 2012), access to EHR is not limited to closed environments. Ac-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

cess has to be complemented with assuring the security of data also. Analysis by Rodrigues et al. (2013b) identifies various confidentiality and privacy issues that should be considered by service providers for moving EHR between EHR-systems.

The constraints in health domain are subject to change with the dynamic creation of demand. It has to handle various contingencies where the system needs to change its behaviour and normal course of actions.

Also, demonstrating access control requirements compliance is difficult and costly due to the level of abstraction required between the requirements and RBAC model. A successful and secured collaboration may exist if the access rules are governed by well-defined authorizations. ABAC implementation can help in predetermining access authorizations. NIST Special Publication 800-53 provides guidance on stringent access controls, for ex. AC-3, AC-16 and many more for securing sensitive information when being processed, transmitted or stored on collaborative systems. The third control i.e. *Mandatory Access Control* (MAC) in the ***Access Enforcement*** family (AC-3 (3)), imposes constraints preventing a subject from passing information to another subject operating at a different sensitivity level, whereas, the fourth control, *Discretionary Access Control* (DAC) in the same family (AC-3 (4)), permits the subject to pass the information to any subject at the same sensitivity level. It is assumed that the problem can be sought through well-integrated access control policies (Hu & Kuhn, 2011) capable enough to dynamically inspect and resolve interoperable issues in sharing of data between heterogeneous systems.

3. DESIGNING OF PROPOSED FRAMEWORK

Lin et al. (2007) algorithm calculates the similarity score between matched and unmatched rules of various access control policies in homogeneous environment. The rules contain various attributes that are individually matched in each rule set. An overall similarity score is then obtained for all the rules belonging to one policy. Similar effort is repeated for all other policies that are probable candidates for collaboration. Finally, the scores are compared and the policies with higher scores are identified to satisfy successful and secured sharing of data by the legitimate users. Our previous work (Bhartiya and Mehrotra, in press) uses this approach to implement and verify the framework in homogeneous healthcare environment. This paper presents the proposed framework to deduce similarities between users in two disparate healthcare organizational hierarchies. Attribute-based access control (ABAC) implementation can reduce the risk of unauthorized access to the data due to its ability of implementing consistent policies and also update them easily to address changing threats. ABAC system is partially dependent on the organization's authorities for assigning attributes related to the allocation of resources (EHR) to the respective subjects (Users). As defined by NIST, "An access control method (ACM) where subject requests to perform operations on objects are granted or denied based on assigned attributes of the subject, assigned attributes of the object, environment conditions, and a set of policies that are specified in terms of those attributes and conditions".

The distribution and composition of ACM must be predetermined to avoid conflicts and other issues weakening the security of sensitive resource. For example, if a resource belonging to one organization is to be shared by other organizations, an unauthorized subject (user) in the organization with lesser restrictions should not have access permissions on it. Consistency in managing, maintaining and employing ACMs is needed to ensure interoperability and comprehensive security.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.1. XACML Environment

Access control framework consistent with ABAC is the Extensible Access Control Markup Language (XACML). The XACML model employs elements such as rules, policies, rule and policy-combining algorithms, attributes (subject, resource, action and environment conditions). XACML architecture includes functions such as Policy Decision Points (PDPs), Policy Enforcement Points (PEPs) and Policy Combining Points (PCPs) to control access. The description of each term is included in Table 1.

ABAC verifies and makes decision on the basis of attribute values assigned under each XACML element in the given rules or policies. Under this arrangement, policies can be created and managed without direct reference to users and resources, and users and resources can be established without reference to policy. In the context of sharing sensitive health information, of the patient, across hospitals, an access control framework (Figure 2) designed on ABAC approach is proposed. Each hospital has a unique categorization of health providers (doctors, nurses and other staff) vertically arranged in the order of their roles and responsibilities. The hierarchies may significantly differ from each other. Hence, these hierarchies need to be standardized in order to generate commonalities between them. A filter-based approach can then be applied to identify the relevant matching roles based on the similarities of hierarchical positions of the users in their respective hospitals. Identifying similarities between user attributes would ensure robust authorization, thereby, enhancing the availability complemented with confidentiality and privacy of health data. Table 2 describes the notations used in the framework.

Table 1. XACML terminologies

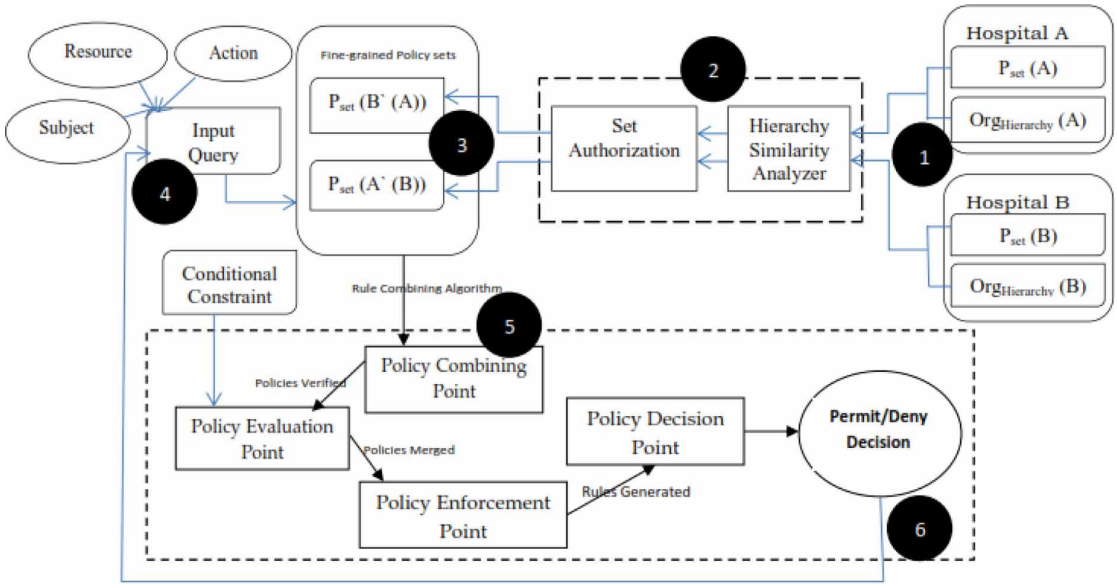
Term	Description
Subject	The entity (user) requesting to access the resource. The subject is sometimes referred to as a requestor
Subject Attributes	Characteristics or attributes of a subject such as name, designation, authority that distinguishes that person from all others
Object/Resource	Entity (data) to be protected from unauthorized use.
Privilege/Action	The authorized behavior of a subject as defined by an authority and embodied in policy or rules.
Environment Condition	Constraints set on the access rules and used as attributes at decision time to influence an access decision.
Policy and rules	A repository of privileges available to subjects from the perspective of the resource that needs protection.
Access control mechanism	Assemble authorization information, which may include information about the resource being protected, the subject requesting access, the policies governing access to the resource, and any contextual information needed to make a decision.
Policy Combining Point (PCP)	To combine disparate policies using rule-combining algorithm
Policy Evaluation Point (PEvP)	To verify the rules against set properties and environmental conditions
Policy Enforcement Point (PEP)	To enforce the decision using SLs and authorizations
Policy Decision Point (PDP)	To render a decision and generate matching rules

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Notations/terminologies for proposed framework

Notations	Description
HSA	Hierarchy Similarity Analyzer generates SL
SL	Security Level determining access rights in policy sets of disparate hospitals
P_{set} (A)	Policy set of hospital A
$Org_{Hierarchy}$ (A)	Organization Hierarchy of hospital A
P_{set} (B` (A))	Fine-grained policy set of hospital A with respect to hospital B
P_{set} (B)	Policy set of hospital B
$Org_{Hierarchy}$ (B)	Organization Hierarchy of hospital B
P_{set} (A` (B))	Fine-grained policy set of hospital B with respect to hospital A

Figure 2. Proposed framework for fine-graining XACML policies using HSA



3.2. Proposed Framework

The framework consists of the following six steps:

Step 1: Identify the policy set and organization hierarchies of the hospitals agreeing to share the data. Compare the position of each user in the given hierarchies and generate similarity scores based on hierarchical distances between them.

Step 2: Input the similarity scores to 'HSA' algorithm for generating unique values to be assigned as Security Level for each user. Input Security Level of each user to 'Set Authorization' algorithm and determine authorization in accordance with other user's security levels.

Step 3: Fine-grain the policy set of hospital A with respect to B and vice versa by assigning security level and authorization with the Subject attribute in applicable rules.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

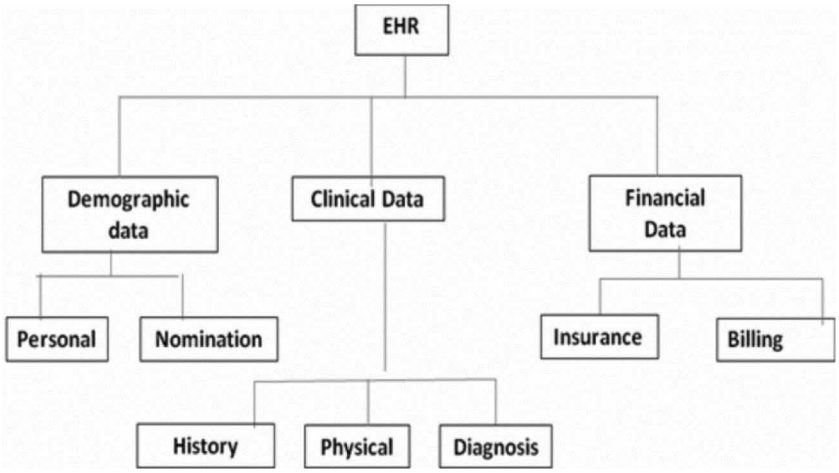
Step 4: Formulate the user request as input query and validate it against the refined policy-sets of hospital A and B.

Step 5: Apply selected rule combining algorithm to compare the rule-sets ***(Policy Combining point)*** and identify matching and non-matching rules based on the input query. Evaluate ***(Policy Evaluation Point)*** and obtain relevant rules that satisfy any pre-set conditional constraints in the policy set(s) of hospital A and B. Merge ***(Policy Enforcement Point)*** the obtained rules and enforce the validated privileges with each rule. Obtain a final decision ***(Policy Decision Point)*** and generate final rule-set that either permit or deny access to the requested data.

Step 6: Notify the decisions obtained to the input query and (if required) repeat the process with new query.

Designing solution to achieve interoperability without compromising security of EHR is highly complex and time-extensive task. This work is a small effort to bridge the interoperable gap and allow authorized access to data in healthcare domain. Heterogeneity in organizational structure and frequent change in user's roles and privileges is experienced in real-time study of healthcare environment. The proposed framework enables to reduce the heterogeneous distances between two structures and ensure authorized and relevant access to the data. Similarity score is calculated for each user attribute in the policy-sets. It lies between 0 and 1. A value near to 1 depicts high similarity between the selected user attributes thus increasing the chances of legitimate and authorized access to the required data. Also, similarity score for unmatched user attribute values is obtained by averaging their score with all the other attributes.

Figure 3. EHR hierarchy in healthcare domain



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. IMPLEMENTATION

4.2. Implementation Environment

Case of two independent healthcare organizations is discussed with respect to sharing of EHRs comprising of various health and personal details of the patient. EHR is categorized and composed in a hierarchy as shown in figure 3. Patient's clinical data comprise of diagnostic, prescriptions and other history details. Demographic data includes patient's personal information i.e. name, address, date of birth etc. It also stores the nominee's detail that is authorized to take decision and handle any contingencies for the patient. Financial data stores the billing and insurance claims of the patient.

4.2.1. Case 1

Environment: Healthcare organization H1 is a full-fledged government-aided hospital comprising of various facilities as per the specialization. It runs an OPD with an inflow of above 2000 patients daily. The user hierarchy for H1 is shown in figure 4(a). MS leads the hospital with each department being managed by respective HOD. A team of Sr. and Jr. Resident doctors handle the OPDs and are assisted by nurses and other technical and administrative staff. Patient is registered by the administrative staff before meeting with the doctor.

Administrative rights and Permissions: The doctor updates the patient's EHR with the history, diagnosis prescription and required laboratory tests details. MS can read record of any patient whereas the HODs can read the records of all the patients treated under their department. Sr. Resident can read and write the clinical details and read demographic details of specified patients. Jr. Resident can read and write the clinical details related to lab test duly approved by the Sr. Resident, but require no permission to write the history details of the patient. Nurses can read demographic and clinical details but have no access to lab details of the patient.

4.2.2. Case 2

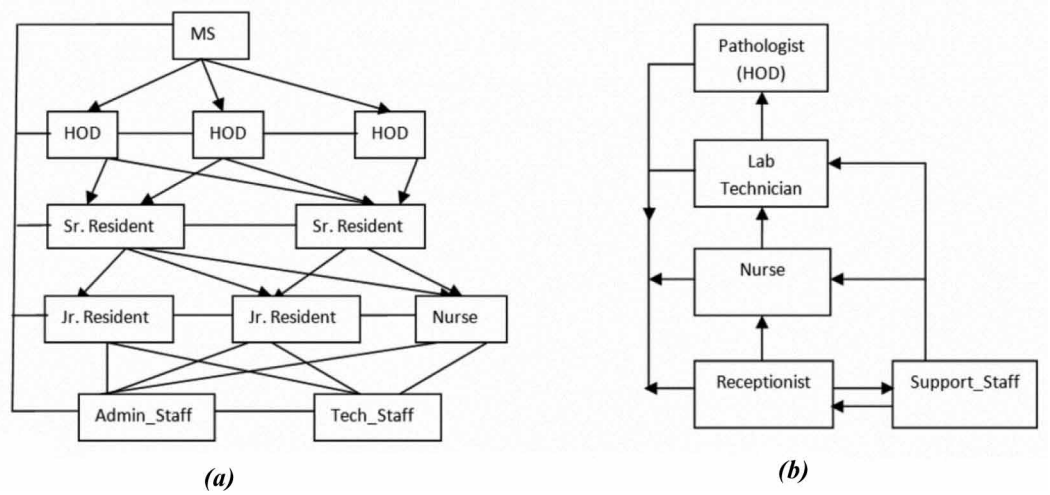
Environment: Another health organization H2 is an independent laboratory conducting various lab tests as prescribed by doctors. It registers various categories of patient i.e. independent or being referred by a doctor or a hospital. The user hierarchy of H2 is shown in figure 4(b). The lab comprise of a Pathologist (HOD), Lab technicians, nurses, receptionist and other administrative staff.

Administrative rights and Permissions: The receptionist records the required test and forwards the patient to the respective room. The nurse checks the basic vitals i.e. the BP, fasting/non-fasting, water intake etc. necessary for conducting the said test. Lab technician performs the test and the report is duly signed by the HOD before dispatching to the patient. The Lab technicians have a read access to the diagnostic details of the patient and write access for updating the test result obtained. The Lab technician need to take patient's or HOD's consent before accessing the diagnostic data. The test results must be verified by the HOD before writing into the patient's EHR. Patient has read permission on his/her test reports. The receptionist and the nurse are restricted to write in the patient's diagnostic section.

Figure 4(a) and 4(b) exhibit the organizational hierarchies of H1 and H2 with respect to users. The position and the responsibilities of HOD, nurses and other technicians vary in both the organizations. Sharing EHRs leads to policy-conflicts thereby resulting over or under exposure of data to the health

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. (a) User hierarchy in health organization H1; (b) User hierarchy in health organization H2



professionals. Thus, a balance between the user domain hierarchies of H1 and H2 needs to be established for secured sharing of the data.

Table 3 reveals the similarity score between the users positioned in these hierarchies (Step 1). Rows represent the similarity scores of H2 whereas columns represent the similarity scores of H1.The scores obtained identifies the similarities between each user-user pair of two different hierarchies. For ex. user hierarchy of H1 shows the HOD as a subset of MS whereas the same has a higher priority in H2. This signifies that some kind of hierarchical similarity exists between these attributes.

Using the hierarchical distances obtained in table 3, HSA assigns SL to each user in the respective category (step 2). The analyser expands the user's association with other users as well as itself. This is done to obtain the unique values as existence of two different security levels for the same association of rules is just not possible. It further enables establishing well-defined authorization on sharing of EHRs in disparate but interoperable healthcare environment. Table 4 shows the SLs assigned to each user thereby ascertaining the level of access permitted to the users specifically while integrating different EHR systems.

A merged set of users are defined for the sake of generality and verification of the proposed framework. The users are grouped into four categories, Doctor (includes Sr. Resident, Pathologist, HOD), Med_tech (includes nurse and Lab technician), admn_tech (includes receptionist, sample collector) and patient.

Table 3. Similarity Scores of user hierarchies in H1and H2

	H1					
H2		HOD	Sr. Resident	Jr. Resident	Nurse	Receptionist
	Pathologist	1	.87	.75	.62	.62
	Lab Tech	.87	1	.87	.75	.75
	Nurse	.75	.87	1	.87	.87
	Sample Collector	.62	.75	.87	1	1
	Receptionist	.62	.75	.87	1	1

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Security_Level (SL) obtained through HSA for each user-pair in H1 and H2

H2 / H1	Pathologist	Lab Tech	Sample Collector	HOD	Sr. Resident	Jr. Resident	Nurse	Receptionist
Pathologist	4	0	0	4	3	2	1	1
Lab Tech	0	4	0	3	4	3	2	2
Nurse	1	2	3	2	3	4	3	3
Sample Collector	0	0	4	1	2	3	4	4
Receptionist	1	2	3	1	2	3	4	4
HOD	4	3	2	4	0	3	2	1
Sr. Resident	3	4	3	0	4	0	3	2
Jr. Resident	2	3	4	0	0	4	4	3

4.2.3. Verification of HSA-Refined Policy Set

Enforcing robust controls and efficiently managing the arrangement of roles and their relationships concludes acceptable level of security mechanisms employed by the organization. The proposed framework need to be verified through properties and constraints as imposed in the real-world scenarios. Verification provides a logical reasoning of the suitability and compatibility of applied access control policies. Verification on merging of access control policies of H1 and H2 must identify and restrict irrelevant and unauthorized exposure of data. First Applicable Rule combining algorithm is executed by the simulator on the defined policies. It produces verification reports where if the property is true in the policy no test cases are generated. On the other hand, if the property is violated, counterexamples are generated.

Figure 5 and 6 enlist a sub-set of Med_Tech access policy defined for H1 and H2 respectively.

Med_Tech of H1 and H2 are permitted write access on Cl_test details. Read access is permitted to the Med_Tech of H2 only. Collaborating H1 and H2 generates permit/deny decisions on the basis of defined properties.

Property: *SPEC (Roles = Med_Tech) & (EHR = Cl_test) & (MLSDefaultAction = write) & (Patient_Consent = True) -> decision = Permit*

Figure 5. Med_Tech access policy rule in H1

```
<!-- ABAC Model: H1-->
    <Rule RuleId="rule_7" Effect="Permit">
     <Target>        <Subject>
        <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">Med_Tech</AttributeValue>
      </Subject>        <Resource>
        <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">Cl_test</AttributeValue>
      </Resource>        <Action>
        <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">write</AttributeValue>
      </Action>      </Target>            <Condition>      <Apply>
        <EnvironmentAttributeDesignator AttributeId="Patient_Consent" DataType="http://www.w3.org/2001/XMLSchema#boolean"/>
        <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#boolean">True</AttributeValue>
     </Apply>      </Condition>      </Rule>
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Med_Tech access policy rules in H2

```
<!-- ABAC Model: H2-->
    <Rule RuleId="rule_6" Effect="Permit">
     <Target>        <Subject>
         <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">Med_Tech</AttributeValue>
        </Subject>          <Resource>
          <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">Cl_test</AttributeValue>
        </Resource>          <Action>
          <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">write</AttributeValue>
        </Action>       </Target>       <Condition>       <Apply>
          <EnvironmentAttributeDesignator AttributeId="Patient_Consent" DataType="http://www.w3.org/2001/XMLSchema#boolean"/>
         <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#boolean">True</AttributeValue>
      </Apply>       </Condition>      </Rule>

    <Rule RuleId="rule_7" Effect="Permit">
     <Target>         <Subject>
         <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">Med_Tech</AttributeValue>
        </Subject>          <Resource>
          <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">Cl_test</AttributeValue>
        </Resource>          <Action>
          <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#string">read</AttributeValue>
        </Action>       </Target>             <Condition>       <Apply>
          <EnvironmentAttributeDesignator AttributeId="Patient_Consent" DataType="http://www.w3.org/2001/XMLSchema#boolean"/>
         <AttributeValue DataType="http://www.w3.org/2001/XMLSchema#boolean">True</AttributeValue>
      </Apply>       </Condition>      </Rule>
```

As mentioned, if the verification result identifies the property to be true in the policy, no test cases are generated. Figure 7 shows the NuSMV verification result where the said property is found to be true in both policies H1 and H2; therefore, no test cases are generated. The decision obtained permits Med_Tech of both the organizations to write the laboratory test details in patient's EHR.

Though Med_Tech, i.e. nurses and lab technicians, are permitted to write the laboratory details in the patient's EHR, it provides no restriction on the amount of data to be accessed by these users. Clinical tests vary from simple, regular health checkups to specific and sensitive tests. Each test report is differently sensitive and needs a security level for defining clear authorization and access permissions. The given policies lack to define authorization and control over the access on the amount of data to be shared.

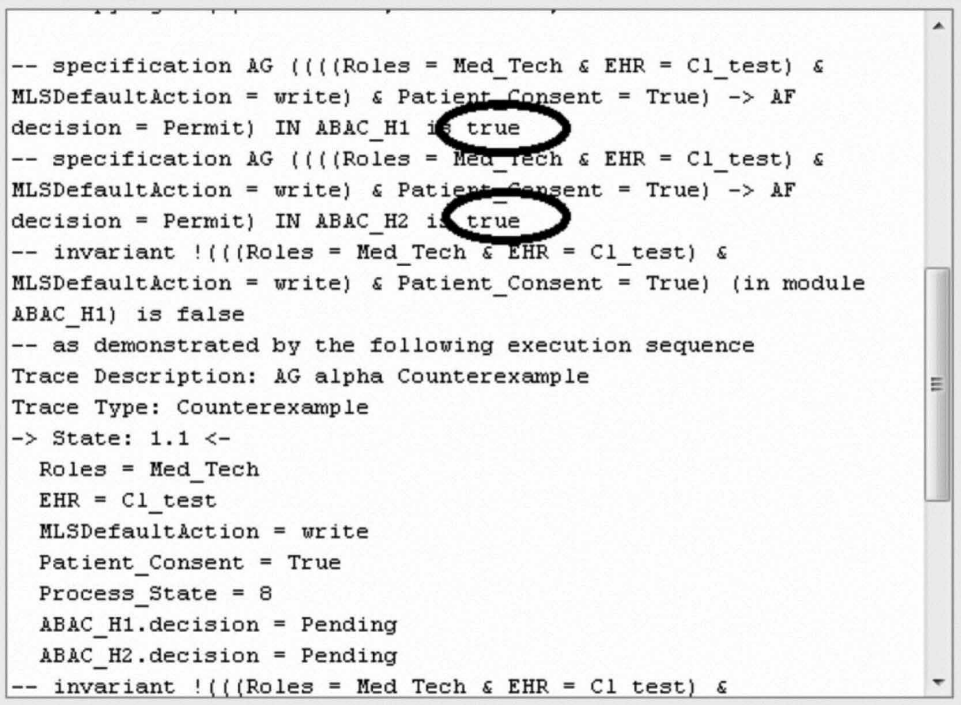
Our framework fine-grains the policies of H1 and H2 by assigning a logistic security level on the given association of two or more users while sharing EHRs. The values in table 4 are assigned as SL to each user-pair of rules defined in H1 and H2 respectively (Step 3). Based on the SL and the data to be shared, the health organization allocates the authorization control on the access of the EHR. Fine-grained policies of H1 and H2 generates permit/deny decisions on the basis of defined properties.

Property: *SPEC (Roles = Med_Tech) & (Security_Level = 2) & (Authorization = Doctor) & (EHR = Cl_test) & (MLSDefaultAction = write) & (Patient_Consent = True) -> decision = Permit*

With respect to the previously defined property, there exist similar set of rules in policies of H1 and H2 as verified in figure 7. The property is redesigned and inputted as request to ACPT (Step 4). Figure 8 shows the NuSMV verification result where the said property is found to be false in H1 and true in H2 (Step 5). The specification identifies the policy conflicts existing in two policies. Though both policies permit a write access on the clinical test details of the patient, differ in context of authorizations and SLs. A counterexample hence generated shows the state of conflict and the possible decision allowing or denying access to the requested data (Step 6).

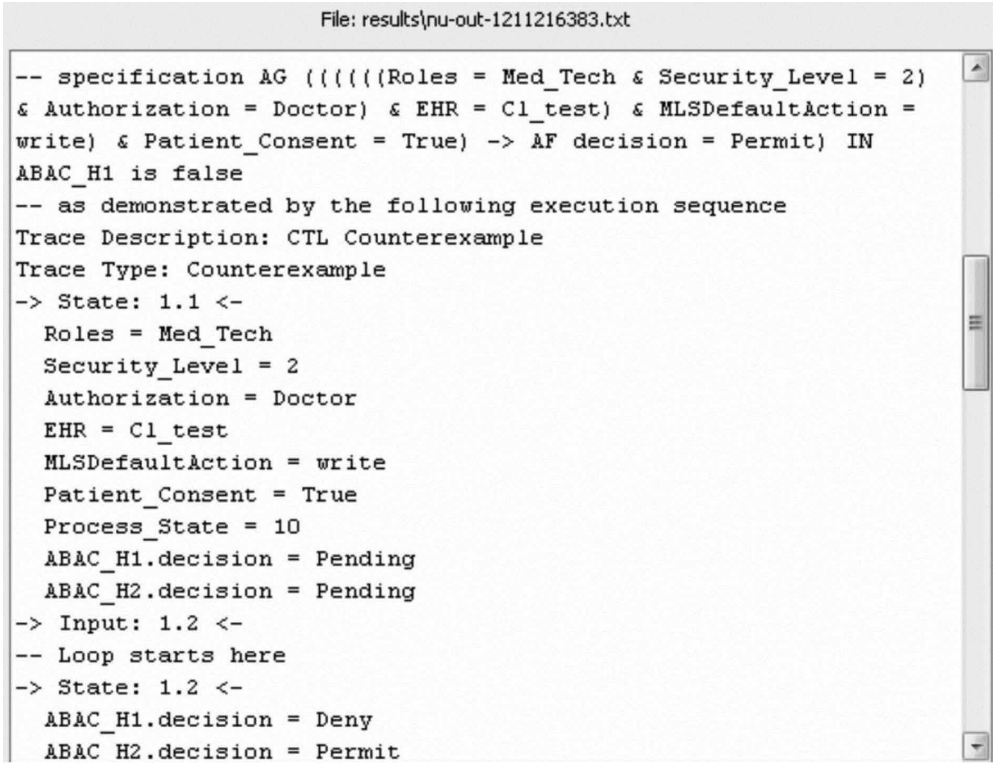
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Property verification result before applying HSA



```
-- specification AG ((((Roles = Med_Tech & EHR = Cl_test) &
MLSDefaultAction = write) & Patient_Consent = True) -> AF
decision = Permit) IN ABAC_H1 is true
-- specification AG ((((Roles = Med_Tech & EHR = Cl_test) &
MLSDefaultAction = write) & Patient_Consent = True) -> AF
decision = Permit) IN ABAC_H2 is true
-- invariant !(((Roles = Med_Tech & EHR = Cl_test) &
MLSDefaultAction = write) & Patient_Consent = True) (in module
ABAC_H1) is false
-- as demonstrated by the following execution sequence
Trace Description: AG alpha Counterexample
Trace Type: Counterexample
-> State: 1.1 <-
  Roles = Med_Tech
  EHR = Cl_test
  MLSDefaultAction = write
  Patient_Consent = True
  Process_State = 8
  ABAC_H1.decision = Pending
  ABAC_H2.decision = Pending
-- invariant !(((Roles = Med_Tech & EHR = Cl_test) &
```

Figure 8. Property verification result after applying HSA



File: results\nu-out-1211216383.txt

```
-- specification AG ((((((Roles = Med_Tech & Security_Level = 2)
& Authorization = Doctor) & EHR = Cl_test) & MLSDefaultAction =
write) & Patient_Consent = True) -> AF decision = Permit) IN
ABAC_H1 is false
-- as demonstrated by the following execution sequence
Trace Description: CTL Counterexample
Trace Type: Counterexample
-> State: 1.1 <-
  Roles = Med_Tech
  Security_Level = 2
  Authorization = Doctor
  EHR = Cl_test
  MLSDefaultAction = write
  Patient_Consent = True
  Process_State = 10
  ABAC_H1.decision = Pending
  ABAC_H2.decision = Pending
-> Input: 1.2 <-
-- Loop starts here
-> State: 1.2 <-
  ABAC_H1.decision = Deny
  ABAC_H2.decision = Permit
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The property verification generates enormous test cases based on m-way factoring. 'm' here denotes to the number of attributes that are logically combined to obtain a valid decision. ACPT allows 2-way, 3-way and 4-way factoring. Table 3 shows a subset of the test cases generated using 3-way factoring to verify the said property on the refined set of access control policies of H1 and H2. The combination of SL, Authorization and Patient Consent is tested for accessing laboratory details by the medical technicians that include nurses and other lab technicians. The rules defined for these users differ in both organizations. For ex. Med_Tech of H2 is permitted a read access on the test details holding SL as 1 and authorization as that of doctor. It clearly contradicts with H1 where no such rule is defined. Another situation exists where the rules are defined in both policies but exhibit disparate combinations of attribute values. For ex. Med_Tech having SL as 2 is permitted an access on the test details if the authorization is of doctor. The rules do exist in H1 and H2 with a variation of patient's authorization instead of doctor's authorization in H1. Hence, the decision obtained reflects a permit in H2 and deny in H1. The verification result identifies such conflicts wherever exist in the rule sets as shown in table 5.

Collaboration of two or more health organizations entails a specific purpose of sharing. An authorization defined on the same set of access rules would vary between disparate collaborations. Our framework exhibits dynamism in assigning the authorizations over the access of data. The framework permits to assign a different authorization on the basis of defined properties even on the similar set of rules in two or more health organizations. Thus, our framework proves to strengthen the security policies in dynamic sharable interoperable healthcare environment.

5. CONCLUSION

Due to different administrative constraints prevalent in heterogeneous systems, achieving interoperability and security while sharing data becomes a challenging task. Each health organization defines an organizational hierarchy and assigns roles and responsibilities to the users on the basis of their qualification and experiences. The users are positioned in a hierarchy ranked according to their ranks and authorities, usually from top to bottom. The ranking and authorization may be differently handled by each organization. Thus, a need arises to determine agreeable access to the data covering the security needs of the

Table 5. Permit/deny decisions obtained on generated test cases to validate the defined property

Role	Resource	Security Level (SL)	Authorization	Patient Consent	Permission	Defined in H1	Defined in H2	Decision H1	Decision H2
Med_Tech	Cl_Test	3	Patient	False	Read	False	False	Deny	Deny
		2	Doctor	True		False	True	Deny	Deny
		2	Patient	True		True	True	Deny	Permit
		1	Doctor	False		False	True	Deny	Permit
		3	Doctor	True	Write	False	False	Deny	Deny
		2	Patient	False		False	False	Deny	Deny
		1	Patient	True		False	True	Deny	Permit
		1	Doctor	True		True	True	Permit	Permit

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

organization and patient as well. To ensure authorized and timely availability of data, a well-integrated secured access control framework is proposed. The framework bridges the interoperability gap and ensures authorized access to the EHRs by balancing the disparities in user's roles and privileges identified in two or more healthcare organizations. Similarity score is calculated to rank the users according to their position in the organization hierarchy. Rule attributes defined in the respective access control policies of disparate organizations are refined by assigning security levels and authorizations to the users. The input query is matched with the refined rule-sets along with the conditional constraints and a justifiable subset of matching rules is obtained. The security level ascertains the sensitivity associated with the data and authorizations ensure only relevant and required access is provided to the users. This further mitigates the problem of gaining additional or full access to the data due to frequent shift of roles and responsibilities of health providers. Another issue very common in heterogeneous environment is the emergence of policy conflicts resulting in either unavailable or unauthorized disclosure of data. The framework is capable in revealing inherent policy-conflicts on merging of disparate policies. The verification tool concludes that such conflicts are traceable and can be resolved. Counter examples generated by the simulator gives an indication of how these conflicts can be resolved.

5.1. Future Scope

Disparities in user's roles and responsibilities have been handled in this work. It can be extended many other disparities categorized into semantic, syntactic, temporal constraints and further devise a solution to mitigate the interoperability and security issues. Apart from resolving identified conflicts, the framework must be able to handle other contingencies for uninterrupted but secured access to the EHRs in interoperable healthcare environment. Timely updating the framework to handling of guidelines issued by various federal agencies like Health Insurance Portability Accountability Act (HIPAA) (Final Rule, 2013) and following the standards such as HL7, DAF (DAF, 2015) can further strengthen its viability and reliability. As stated by NIST, dual authorization, i.e. approval of two authorized individuals is mandatory for permitting the request can be incorporated with very little effort in the current framework.

REFERENCES

Amini, M, & Jalili, R. (2010). Multi-level authorization model and framework for distributed semantic-aware environments. *Information Security*, 4(4), 301-321.

Bai, Y., Dai, L. & Li, J. (2014). Issues and Challenges in Securing eHealth Systems. *International Journal of E-Health and Medical Communications*, 5(4), 1-19.

Bhartiya, S. & Mehrotra, D. (2014). Challenges and Recommendations to Healthcare Data Exchange in an Interoperable Environment. *Journal of Health Informatics*, 8(2), e16.

Bhartiya, S., & Mehrotra, D. (in press). An Access Control Framework for Secured Sharing of Electronic Health Records using Hierarchy Similarity Analyzer. *Int. J. of Electronic Healthcare*.

Data Access Framework (DAF) Charter and Members. (2015). Project Charter, Creative Commons Attribution Share-Alike 3.0 License. Retrieved from http://wiki.siframework.org/Data+Access+Framework+Charter+and+Members

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

de la Torre-Diez, I., López-Coronado, M & Rodrigues, JJ. (2012). How to Measure the QoS of a Web-based EHRs System: Development of an Instrument. *Journal of Medical Systems*, 36(6), 3725-3731.

Fernández-Cardeñosa, G., de la Torre-Diez, I., López-Coronado, M. & Rodrigues, JJ. (2012). Analysis of Cloud-based Solutions on EHRs Systems in Different Scenarios. *Journal of Medical Systems*, 36(6), 3777-3782.

Rule, F. (2013). Modifications to the HIPAA Rules, Rules and Regulations, Department of Health and Human Services, Office of the Secretary, 45 CFR Parts 160 and 164. *Federal Register*, *78*(17).

Health Level Seven International. (2015).HL7 Standards - Section 1: Primary Standards. Retrieved from http://www.hl7.org/implement/standards/product_section.cfm?section=1

Health Level Seven International. (2015). HL7 Fast Healthcare Interoperability Resources Specification (FHIR) Release 1. Retrieved from http://www.hl7.org/implement/standards/product_brief.cfm?product_id=343

HL7 Evolution. (2010). Corepoint Health. Retrieved from https://www.corepointhealth.com/sites/default/.../hl7-v2-v3-evolution.pdf

Hu, V. C., & Kuhn, D. R. (2011). *Model Checking For Verification Of Mandatory Access Control Models And Properties. International Journal of Software Engineering and Knowledge Engineering*, *21*(1), 103–127.

Hu, V., Schnitzer, A., & Sandlin, K. (2013). *NIST Special Publication 800-162*. Attribute Based Access Control Definition and Considerations.

Huang, F. Z., & Liu, L. (2009). A DL-based method for access control policy conflict detecting. *Proceedings of the First Asia-Pacific Symposium on Internetware*.

Hwang, J., Xie, T., Hu, V., & Altunay, M. (2010). ACPT: A Tool for Modeling and Verifying Access Control Policies. *Proceedings of the 2010 IEEE International Symposium on Policies for Distributed Systems and Networks POLICY '10*, Washington, DC, USA (pp. 40-43). 10.1109/POLICY.2010.22

Imine, A., Cherif, A., & Rusinowitch, M. (2009). A Flexible Access Control Model for Distributed Collaborative. In W. Jonker and M. Petkovic (Eds.), *Secure Data Management, LNCS* (*Vol. 5776,* pp. 89–106). 10.1007/978-3-642-04219-5_6

Karp, A. H., Haury, H., & Davis, M. H. (2010). From ABAC to ZBAC: The Evolution of Access Control Models. *Proceedings of the International Conference on Information Warfare* (pp. 202-211).

Kuhn, D. R., Hu, V., Ferraiolo, D. F., Kacker, R. N., & Lei, Y. (2015). Pseudo--exhaustive Testing of Attribute Based Access Control Rules. *Proceedings of the Intl. Workshop on Combinatorial Testing.* Retrieved from http://csrc.nist.gov/groups/SNS/acts/documents/abac-pseudo-ex-iwct.pdf

Li, A., Li, Q., Hu, V. C., & Di, J. (2015). Evaluating the Capability and Performance of Access Control Policy Verification Tools. Proceedings of the IEEE Military Communications Conference MILCOM '15. Retrieved from csrc.nist.gov/groups/SNS/acpt/documents/milcom15_final.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lin, D., Rao, P., Bertino, E., & Lobo, J. (2007). An Approach to Evaluate Policy Similarity. *Proceedings of the 12th ACM symposium on Access control models and technologies* (pp. 1-10). 10.1145/1266840.1266842

Miyagawa, S., Yamasaki, S., Uchiyama, E., & Amoroso, D. L. (2014). Framework for Information Sharing with Privacy and Priority Control in Long-Term Care in Japan. *International Journal of E-Health and Medical Communications*, *5*(1), 46–62. doi:10.4018/ijehmc.2014010103

Park, J. H., & Lee, D. G. (2007). PIS-CC RBAC: Patient information service based on CC_RBAC in next generation hospital considering ubiquitous intelligent environment. *Proceedings of the International Conference on Multimedia and Ubiquitous Engineering (MUE'07)* (pp. 196-200). 10.1109/MUE.2007.171

Parker, S., & Wisely, S. (2009). *Guide to Information Sharing and Data Interoperability for Local Communication Centers in IJIS Institute, Public Safety Data Interoperability Program (PSDI)*. U.S. Department of Justice Office of Justice Programs.

Power, D., Slaymaker, M, & Simpson, A. (2011). Conformance checking of dynamic access control policies. Proceedings of the *13th International Conference on Formal Engineering Methods, ICFEM 2011*, (Vol. 6991*,* pp. 227-242) Durham, UK, October 26-28, 2011.

Rodrigues, J. J. P. C., de la Torre-Doez, I., Fernández Cardeñosa, G., & López-Coronado, M. (2013a). Security and Privacy Requisites Analysis of Cloud-based EHRs Systems [JMIR]. *Journal of Medical Internet Research*, *15*(8). PMID:23965254

Rodrigues, J. J., Pedro, L. M., Vardasca, T., de la Torre-Díez, I., & Martins, H. M. (2013b).. . *Mobile Health Platform for Pressure Ulcer Monitoring with EHR Integration, Health Informatics Journal, SAGE Publications*, *19*(4), 300–311.

Security and Privacy Controls for Federal Information Systems and Organizations. (2013, April). National Institute of Standards and Technology Special Publication 800-53, Revision 4. Retrieved from doi:10.6028/NIST.SP.800-53r4

Thomas, R. K. (2007). Team-based access control (TMAC): a primitive for applying role-based access controls in collaborative environment in RBAC '97. *Proceedings of the second ACM workshop on Role-based access control* (pp. 13–19)

Wickramasuriya, J., & Venkatasubramanian, N. (2004). Dynamic Access Control for Ubiquitous Environments, In R. Meersman, Z. Tari (Eds.), In *On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE* (pp. 1626–1643)

This research was previously published in the International Journal of E-Health and Medical Communications (IJEHMC), 7(2); edited by Joel J.P.C. Rodrigues; pages 14-29; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 75
Patient Data De-Identification: A Conditional Random-Field-Based Supervised Approach

Shweta Yadav
Indian Institute of Technology Patna, India

Asif Ekbal
Indian Institute of Technology Patna, India

Sriparna Saha
Indian Institute of Technology Patna, India

Parth S Pathak
ezDI, LLC, India

Pushpak Bhattacharyya
Indian Institute of Technology Patna, India

ABSTRACT

With the rapid increment in the clinical text, de-identification of patient Protected Health Information (PHI) has drawn significant attention in recent past. This aims for automatic identification and removal of the patient Protected Health Information from medical records. This paper proposes a supervised machine learning technique for solving the problem of patient data de- identification. In the current paper, we provide an insight into the de-identification task, its major challenges, techniques to address challenges, detailed analysis of the results and direction of future improvement. We extract several features by studying the properties of the datasets and the domain. We build our model based on the 2014 i2b2 (Informatics for Integrating Biology to the Bedside) de-identification challenge. Experiments show that the proposed system is highly accurate in de-identification of the medical records. The system achieves the final recall, precision and F-score of 95.69%, 99.31%, and 97.46%, respectively.

DOI: 10.4018/978-1-7998-1204-3.ch075

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

With the start of the golden era in the medical interpretation, the vast amount of information in the clinical domain is increasing at a rapid rate. In the past decade, with the development of the health information technology and health data documentation, there has been progress in how heath care is performed (Berner et al., 2005).

With the widespread use of health information technology, there has been huge pace in the increment of clinical data in addition to the fast adoption of the Electronic Clinical Records and with the conversion of narrative data to the electronic form. The amount of information can be improved further with the minimization of the medical error. This requires the development of some sophisticated tools for Medical Language Processing (MLP). Most medical records are in the narrative forms which are formed as the result of transcription of dictations, direct entry by providers, or use of speech recognition applications. However, their use in this form is restricted to any organization or research, as medical records have a sufficient number of personal health information or protected health information (PHI). According to Health Insurance Portability and Accountability Act (HIPAA), 1996, the PHI terms need to be enclosed and protected. This has lead to de- identification problem. Paragraph 164.514 of the Administrative Simplification Regulations promulgated under the Health Insurance Portability and Accountability Act (HIPAA) states that for data to be treated as de-identified, it must clear one of two hurdles (HIPPA ACT 1996).

1. An expert must determine and document "that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information."
2. Or, the data must be purged from a specified list of seventeen categories of possible identifiers relating to the patient or relatives, household members and employers, and any other information that may make it possible to identify the individual.

Studies showed that there was a significant drop in the patient consent request reducing the participation rate and also, this is quite infeasible for the huge population. Even, in the case when a patient provides the permission, documents must be tracked to stop any unauthorized disclosure. This emerging problem of consent, waiver, and tracking can be effectively handled if the patient personal health information is properly de-identified facilitating the clinical NLP research (Wolf & Bennett, 2006).

De-identification task is more specifically defined as the step where the private information is removed or replaced while keeping the record as it is (Stubbs et al., 2015). De-identification is a type of traditional named entity recognition (NER) problem, with the property of defining a term to be PHI type or not. The main aim of de-identification challenge as pointed out earlier is to remove the PHI terms maintaining data integrity as much as possible. Every record is enclosed in the RECORD_ tags and is provided a unique ID which is randomly generated. Figure 1 shows Sample Discharge Summary Excerpt; a sample discharge summary from the training dataset where the goal is to identify the PHI (private health information) terms. In this summary, some the PHI terms are doctors' name ("Dr. Do Little"), patient name ("John Doe") and hospital name ("ABHG", "SBHG"). A TEXT_ tag encloses the text of different records. Each PHI instance is enclosed within PHI_ tags and the PHI TYPE represents the category of the PHI term as shown in Figure 1.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Sample discharge summary excerpt

```
<RECORD ID="641">
<TEXT>
<PHI TYPE="HOSPITAL">SBHG</PHI>
HISTORY OF PRESENT ILLNESS:
Mr. <PHI TYPE="PATIENT">Doe</PHI> is a XX-year-old white male with a history of
diabetes mellitus , inferior myocardial infarction , who underwent open repair of his
increased diverticulum <PHI TYPE="DATE">January 1st</PHI> at <PHI
TYPE="HOSPITAL"> ABHG</PHI> .
No cardiopulmonary resuscitation was performed as per the patient and family wishes
DISCHARGE SUMMARY
ARF32 FA
DISCHARGE SUMMARY NAME:
UNIT NUMBER:
<PHI TYPE="PATIENT">JOHN DOE</PHI>
ADMISSION DATE:
<PHI TYPE="DATE">01/01/1900 </PHI>
DISCHARGE DATE:
<PHI TYPE="DATE">01/02/1900 </PHI>
DD:
<PHI TYPE="DOCTOR">Dr. Do little</PHI>, M.D.
</RECORD>
```

As shown in Figure 1, the task is to enclose the PHI terms such as (“Dr. Do Little”, “John Doe”, “ABHG”, “SBHG”). This task can be seen as the typical sequence labeling task where for e.g. “Dr. Do Little” should be labeled as whole with Doctor. Here, we present an example where input forms the sequence of words and output is the label sequence and its corresponding de-identified sentence. “BIO” notation was followed to label the NE where “B” represents the beginning of label sequence, “I” denotes intermediate of label sequence and “O” represents others.

In the past few years, de-identification of the clinical record has become the booming research area. Sixth Message Understanding Conference (MUC) had defined the Named Entity Recognition (NER) problem with an objective to identify the relevant entities from the text. NER forms an important task in NLP. Similarly, in the medical domain, NER is very crucial in the identification of the clinical terms by which we can get insights to the problem, medication, procedure, and examination.

Major research works for entity extraction in the medical domain had used machine learning and rule-based concepts. MedLEE (Friedman et al., 1996) is one such system in clinical domain proposed by Carol Friedman et al. at Columbia University. Aronson (2001) developed MetaMap system at National Library of Medicine. Another system cTAKES (Savova et al., 2010), developed using Unstructured Information Management Architecture (UIMA) framework and OpenNLP, a widely accepted system.

Recently, the Medical Natural Language Processing Community had organized a number of challenges for NER task. The earliest shared task is organized by Center of Informatics for Integrating Biology and the Bedside (i2b2) in 2009 on medical name recognition (Uzuner et al., 2010). In 2010, i2b2 shared task was conducted with the aim to identify the medical problems, test entities and various treatments from

Table 1. Example of PDI task following BIO notation

Sentence	Mr.	John	D	Doe	is	xx-year-old	white	male
Named Identity	O	B-Doctor	I-Doctor	I-Doctor	O	O	O	O
De-Identified Sentence	Mr.	XYZ_Doctor			is	xx-year-old	White	male

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the clinical records (Uzuner et al., 2011). Since then, Conference and Labs of the Evaluation Forum (CLEF) had organized a challenge in 2013 with the aim of solving disorder mention recognition and normalization (Goeuriot et al., 2013). Very recently, i2b2 in 2014 (Stubbs et al., 2015) had organized a challenge for patient data de-identification problem and CUI (concept unique identifier) detection. Semantic Evaluation (SemEval) 2014 also came up with the disorder mention recognition and the normalization. With a large number of challenges organized almost every year, it had drawn the attention of many researchers to solve these real-life challenging problems.

Different approaches have been used to solve this problem ranging from machine learning techniques (Stubbs et al., 2015; Yang et al., 2015), rule-based system (Beckwith et al., 2006; Berman et al., 2003) to the hybrid system. Rule-based systems are designed over dictionaries with the rules that are designed manually to recognize the PHI terms from the text. Building a system using rules lacks generalization capability as well as it consumes more time and skill for developing rules. Rule-based systems are seen to provide good performance for the PHI terms those appear less often. This has lead to the development of some machine learning techniques for automatic extraction of PHI terms from the text which makes the process more generalized. The only drawback of machine learning based approach is the requirement of having large annotated corpus which can be developed manually.

Generally, the NER problem is the sequence labeling problem where the aim is to identify the proper output sequence of the entities. Thus, for the given input sequence of words, we need to obtain the best labeled- sequence. Researchers have primarily used supervised machine learning algorithms to solve this kind of medical NER problem. The algorithms used for extracting medical entities are Conditional Random Fields (CRFs) (Lafferty et al., 2001), Maximum Entropy (ME) (McCallum et al., 2000), and Structural Support Vector Machines (SSVMs) (Xue et al., 2008). Among these different classifiers, Conditional Random Field had drawn a huge attention for solving the sequence labeling task. Top best system (Yang et al., 2015) of de-identification task had used CRF as the underlying classifier because of the property of CRF to properly map the relationships between their surrounding tokens. Supervised machine learning techniques are able to capture the handcrafted features very effectively. Some of the prominent features which have been used are orthographic features that capture the word structures, syntactic features like Part of Speech (PoS), and semantic features as well as the discourse features. These features are typically used for solving any NER problems. Many other domain specific features could be generated for further improving the performance of the system.

In the current work, we develop a supervised machine learning based classification framework for solving the patient de-identification problem. The proposed method is evaluated on the i2b2 2014 dataset. It concerned identifying four different PHI categories, namely PATIENT, DOCTOR, HOSPITAL names and DATE from the medical records. Taking into account the effectiveness of applying CRF in solving the NER problem, CRF was adopted to train the model. Diverse features are extracted and CRF based model was trained using this huge collection of features. Several experiments were carried out with different cross folds. Results of the proposed system are shown for different cross-validation settings like 5,6,7,8,9,10 fold cross validations.

This task suffers from various challenges as similar to any NLP task. One such challenge is the ambiguity problem where we found that due to the lexical similarity PHI terms mostly overlap with the non-PHI terms. An example includes Brown (Doctor name) which is a PHI term vs. brown (Colour name) which is a non-PHI term. Another major challenge is the occurrence of some out-of-vocabulary PHI terms which are not present in the dictionaries. These words act like the foreign words which can be misspelled. Gazetteers list, dictionaries are the mostly used for solving any NER problem but the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

contribution of these types of words lowers the performance and creates some errors. Results obtained prove that the proposed system outperforms the existing systems in identifying four different PHI terms.

The remaining paper is organized as follows. Next immediate section discusses the background of de-identification task and its significance. This is followed by a section which provides the insight to the related work. Next section discussed the methodology adopted for the de-identification which is followed by the details of experiments and their analysis. Finally, we conclude in the last section.

BACKGROUND AND SIGNIFICANCE

Most of the works done in biomedical Natural Language Processing (NLP) domain can be distinguished into two different types on the basis of the text, biomedical text and clinical text (Meystre et al., 2010). Biomedical texts are the texts that appear in the journals, articles, posters etc. On the other hand, the clinical texts are mostly the texts written by the doctors in the form of prescriptions for the patients. This text comprises of the patient information like his/her name, age, gender, the procedure that was applied to him/her and what are his/her medical histories. These texts also provide insights to the findings made during procedure and the medication provided to them.

The clinical text is the whole range of narrative found in the patient records. These texts vary from the short-segment which gives the chief complaints to the long-text comprising of the history of the patients. There is another class of text which is discussed in the literature very often (Coorevits et al., 2013). While some of the texts are similar to that of biomedical texts for example internal research reports, the others appear to be like clinical texts for example patient notes. As these narratives are rarely exposed outside the organizations, there is a very sparse study on them.

There is one obvious question that what distinguishes clinical text from the biomedical text and why processing of these poses a challenge to the NLP society. Firstly, in the clinical texts, a large number of grammatical errors can be found which may have occurred at the telegraphic phrase. Other segments in clinical documents which create some problem are discharged summaries and the lab test reports. Moreover, clinical documents suffer from the problems like abbreviations and acronyms. These lexical structures are often overloaded i.e., there is a multiple rendering for the same word set. Literature shows that almost 33% of the times, acronyms are overloaded which makes it prone to the ambiguity problem (Liu et al., 2001). There is a huge problem due to the spelling mistake, which is observed most commonly in the notes lacking spelling support. US Veterans Administration (VA) has one of the largest EHR systems with only the basic text support. Even in their corpus, there are lots of abbreviations and acronym problem. Furthermore, the clinical data consists of different types of characters which appear anywhere within the text. Another problem that is raised is to change the unstructured narrative text into the structured; templates along with pseudo-tables are in general used in the method of conversion (Meystre et al., 2010). Template having a narrative structure for patient history or the discharge summary that is common among the medical society provides benefits to NLP. Explicit templates are formatted text with the fields specific to the institution are filled by the user. These different variations in template make this domain quite challenging.

There are many methodologies adopted to cover different types of questions related to clinical research domain (Dehghan et al., 2015), (Liu et al., 2015), (He et al., 2015). Although, the pharmaceutical industry covers the clinical trails and therapy but nowadays, these industries are spreading the horizon of solving the medication issues like drug interaction, disease prevention through the proper investigation of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

diseases. These include the course of the disease and various other factors like, what were the different criteria for diagnosis, how the patient education affected his/her health and his/her surveillance. With the rapid and booming research interest in the clinical domain, the studies have also spread to obtain the role of different gene pathways in the identification of their relations in the development of diseases. Some clinical research focuses on improving the health care with lowering the health care cost, thus providing the benefits to the mass. As such, it is highly important to have a proper study of the clinical records in solving the major unsolved and challenging problems of the society with the NLP system.

De-identification task forms one such challenge which aims to identify the noun phrases or part of a noun phrase automatically from the text. In the traditional named entity recognition task, originated by the MUC challenge, the focus was to identify the person, organization, location names from the newspaper type electronic text. Generally, de-identification task is the task of identifying and enclosing patient personal health information like not just identifying the persons but also identifying the type of person (Patient, Doctor), Organizations (Hospital), Location (Streets, Country), ages, dates but not in the term of years. Apart from these, structures of the text too vary a lot. While in newswire or journalistic domain, lots of clues can be obtained from their titles, abstract and cases while in the patient records, due to its semi-structured form the task becomes quite challenging.

LITERATURE SURVEY

De-identification task is a type of the traditional Named Entity Recognition (NER) problem. Message Understanding Conference had been the center for the evaluation of NER system in several domains. MUC had organized several challenges to identify several named entities like person, location, organization, time, dates etc. For the MUC-6 NER task, a total of 15 teams participated in the challenge submitting 20 different systems. The best F-Measure value reported was 96.42% (Sundheim, 1997). In the MUC-7 challenge, 14 systems were submitted from 12 different teams obtaining the highest F-Measure value of 93.39% (Chinchor, 2001). National Institute of Standards and Technology (NIST) organized two different tasks, Information Extraction Named Entity task and the Automatic Content Extraction (ACE) tasks. Both of these tasks were evaluated on the newswire domain.

In the biomedical domain, Critical Assessment of Information Extraction Systems in Biology (BioCreAtIvE) (Hirschman et al.,2005), organized several biomedical challenges including the entity extraction task, where the task is to identify different entities like gene name, protein name from the given biomedical text derived from the PubMed, PMC. PubMed and PMC are huge repositories consisting of almost 3.5 million biomedical literature. For the evaluation, TREC Genomics Track was considered. The BioNLP 2009 challenge (Kim et al., 2009) focuses on event extraction primarily on the protein and gene event extractions. The same challenge was also focussed on BioNLP 11(Pyysalo et al., 2011) and BioNLP 13 shared task (Nedellec et al., 2013), where data was used from MEDLINE. Another task BioASQ was organized for entity extraction using data from PubMed. Several challenges have been organized for entity extraction. Recently, i2b2 had organized the task for the de-identification of the clinical terms in the years 2006 and 2014 (Stubbs et al., 2015). Here, the challenge was to identify the certain PHI terms.

In the 2006 shared task, a total of 889 patient records were used. This challenge got sixteen submissions from seven different teams. The challenge had used longitudinal records of each patient rather than the individual records. It was observed that longitudinal records contain more personal information of patients rather than an individual record. Completely identifying the patient information from individual

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

records is a problematic task. This makes the de-identification using the longitudinal records a more complicated task. The submitted systems identified the task as a token classification task (using Support Vector Machine) while some distinguished it as the sequence labeling task (using Hidden Markov Model, Conditional Random Field).

Wellner et al. (2007) in i2b2-2006 challenge built the best system based on CRF. They used some lexical and semantics features. They adapted two important NER systems, one based on the CRF, Carafe, and other Lingpipe, HMM-based implementation. The carafe was seen to be highly effective in capturing the PHI terms. Szarvas et al. (2007) adopted an iterative method for named-entity de-identification. They treat de-identification as a classification task and use machine learning based decision tree with local features and dictionaries. Aramaki et al. (2006) also used CRF as their classifier with the features such as local, external and global features. The global feature includes the features which capture position, the length of a sentence and also several tokens in the previous sentence. External features include the dictionary for location, person name, and dates. They performed two-phase learning wherein the first phase, they learned their system using local features and in the second phase, they include first phase features set in addition to the most frequent label. They identified that sentence feature is very informative in identifying dates, IDs and patient because of the appearance of this PHI in beginning of the records while the external feature was useful in identifying dates, locations, doctors and patients.

While many systems have used CRF based classifier, some submitted systems also used SVM. Hara (2006) developed a hybrid system making use of rules. The rules identified the patterns for the PHI terms such as phone numbers, IDs, and dates. They train SVM on the global and local features which were found to be very helpful in identifying the PHI terms such as a hospital, patient, location, doctors, and age. In addition to these features, their system included the section heading which is closed to the target, category of the sentence, the surface form of target, the root, PoS tag and orthography features. The system used two types of the ordered tree, dependency tree and "n-grams trees" making use of the Boosting Algorithm for Classification Trees (BACT). With this information and in addition of chunking technique, they identified the position and the text span of PHI terms such as a patient, doctor, location, and hospital. The results show that performance is degraded with the incorporation of sentence feature.

In 2014 patient de-identification task, machine learning technique was more prominent. According to the 2014 i2b2/UTHealth shared task, the challenge was to de-identify the data in a stricter way as compared to the HIPAA standards in the form of additional categories. The task was to identify the 25 PHI categories with the associated subcategories from the clinical records like professions, time and dates, facilities etc. These tasks also consider the longitudinal records. Yang et al. (2015) submitted the system in the i2b2-2014 challenge and they reported the accuracy of 93.60%. This was the best performing system. They developed a hybrid model which used both the machine learning and the rule-based technique to extract the PHI terms. They exploited the various features like linguistic features, syntactic feature and different word surface oriented features which were further improved by using the task-specific features. They also used the regular expression pattern to identify the PHI terms like date, ID.

The second best system (Dehghan et al., 2015) from Manchester used a combined system developed using the knowledge-based technique as well as the rule-based system. The rules were defined to obtain the orthographic features, different patterns. They used CRF as the underlined classifier to obtain the PHI terms. For some categories like PATIENT, DOCTOR, HOSPITAL, the initial annotation was extracted at the patient level and dictionary was created for the patient. At the last stage, the outputs of the rule-based system, as well as the CRF based model, were integrated on the basis of some rules. These systems were highly efficient in obtaining the PHI terms such as the DATE, DOCTORS, and PATIENT.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

System (Liu et al., 2015) from Harbin Institute of Technology Shenzhen Graduate School in i2b2 challenge used two different variations of the CRF. They define one CRF system working at the token level while another CRF system that works at the character level. They performed feature engineering and used the features such as bag-of-words, section relevant features, lexical features, semantic features. Apart from the two different CRF outputs, the third output was obtained by developing the rules using the Rule-based system. The final outputs of all the three systems were merged on the basis of the output overlapping. If there was no overlapping i.e., if all the three systems obtained the same PHI tag, then that tag was accepted else the tag was obtained on the basis of the preference. They give the first preference to the rule based system, then the character-level system and lastly to the token-level based CRF system.

Another submitted system from Harbin Institute of Technology (He et al., 2015) used OpenNLP system to pre-process their data, sentence splitting, and tokenizer. They trained their system using CRF making use of features such as lexical, syntactic and orthographic features. They did not use external features such as dictionaries to identify the terms. The system developed by Kaiser Permanente (Torii et al., 2014) adapted the MIST tool. They include some additional rules that used lexicon for Location and Profession and different regular expressions for the categories such as Phone Numbers, ZIP and Organization. In addition to MIST system, the team learned NER model using Stanford NER system. The system then merges the overlapped identified entities. They reported the F-Measure value of 92.21%.

Through the literature survey, we found that hybrid systems (both the rule-based system as well as the machine learning based) perform well when compared with the individual techniques. It is reported that among top ten ranked systems, seven systems have adopted the rule-based technique. MedEx (Xu et al., 2010) forms one such system proposed at Vanderbilt University that utilizes the biomedical dictionary to identify the important medical concepts.

Overall, the best systems report high recall and precision, often greater than 90%, and sometimes as high as 99%. Nevertheless, no study has evaluated the performance of automated de-identification for all PHI classes (Meystre et al., 2010). Only a handful of studies provide details on over-scrubbing (non-PHI wrongly identified as PHI) and none of them investigate the effect of de-identification on subsequent IE tasks. It is indeed possible that de-identification has an adverse effect on IE accuracy (Meystre et al., 2010). Over-scrubbing errors could overlap with useful information, for example, if a disease name is erroneously recognized as a person name, it will be removed and lost to subsequent IE applications. Second, NLP techniques such as part-of-speech tagging and parsing may be less effective on modified texts.

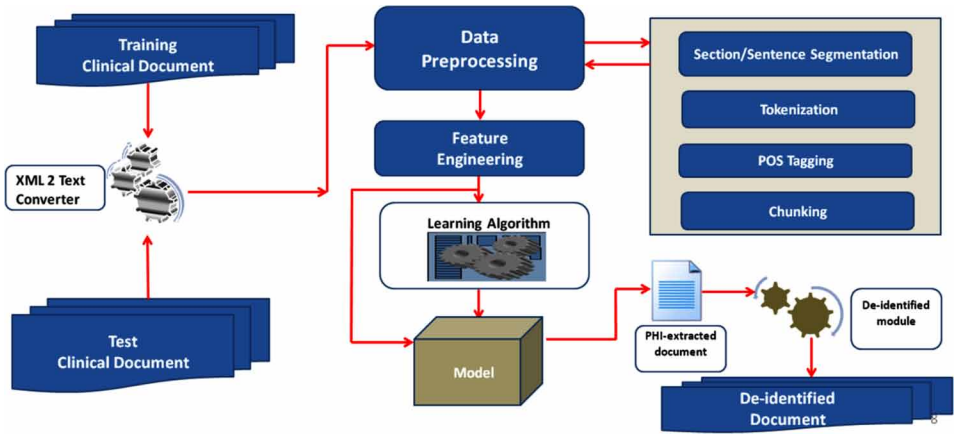
It can be concluded that though a lot of systems have evolved in the recent past, but still there is a need to develop a more accurate automated de-identification system.

METHODS

Supervised machine learning based classification frameworks have been developed to detect the PHI tokens from the medical records. Similar to any entity extraction technique, de-identification task is also considered to be the sequence labeling task. The proposed system architecture is divided into four different modules. The first module deals with the pre-processing of data. The second module represents the feature engineering part where features are generated. Training of the model through Conditional Random Field (CRF) is done in the third module. The fourth module represents the de-identification of the PHI terms. Detailed descriptions of the methodology adopted for the system are given below. Figure 2 represents the Proposed System Architecture.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Proposed system architecture



Pre-Processing of the Medical Records

Datasets provided with the i2b2 de-identification challenge were untokenized; this necessitates the use of the pre-processing module for the process of tokenization. The pre-processing was performed in three different stages. The first stage of the pre-processing module deals with the sentence splitting. The second module focuses on the tokenization of the medical records. The last and final stages are used to identify the rules for tokens such that when they are divided, they should match some patterns.

Apart from this process, pre-processing steps also include shallow parsing and Part of Speech (PoS) tagging to obtain the word lemma, chunk and part-of-speech of the corresponding words. Additionally, we extracted some of the document- level features, for example obtaining the different sections (HISTORY OF PATIENT ILLNESS, MEDICATIONS etc.) and position of the sentence in the text (beginning of the text, end of the text) using some of the hand-crafted rules.

Feature Generation

This module focuses on developing some prominent features. The features were developed, such as the linguistic, syntactic features as well as word surface-oriented features in order to capture the semantics of the PHI terms. The features were further enhanced by embedding the feature set by the task-specific features. The list of the extracted features for the PHI recognition task is given below:

- **Token Features:** Some features are generated during the pre-processing step like word lemma, word chunk and the Part-of-Speech tag of the current word.
- **Bag-of-Word Feature:** It includes the unigrams, bigrams, trigrams of the target token within the window size of [-2, 2]. An *n*-gram is a contiguous sequence of *n* items from a given sequence of text or speech. An *n*-gram of size 1 is referred to as an unigram, size 2 is a bigram and size 3 is a trigram.
- **Contextual Features:** This feature is the combined feature which includes the lemma, chunk, POS tag of the surrounding tokens. We consider the window size of [-3, 3]. These features give the context knowledge for each word.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Combined POS-Token Feature:** The combination of POS tag within the context of [-1, 1] and the current token $\{w_0p_{-1}, w_0p_0, w_0p_1\}$ is formed, where w_0 represents the current word, and p_{-1}, p_0 and p_1 represent the previous POS tag, current POS tag and the next POS tags respectively.
- **Affixes:** This feature includes prefix and suffix of the word. It refers to the fixed length characters stripped either from the left most part or the rightmost part of the token. For example, the suffix and prefix of word 'health' are h, he, hea, heal and h, th, lth, alth respectively. Here, we set the length to size 4.
- **Sentence Information Feature:** This feature includes the structure of the sentence. It includes the information regarding length of the sentence, whether the sentence has any delimiters like '.', '?', '!', whether there is any bracket unmatched.
- **Orthographic Feature:** Several orthographic features were developed to capture the capitalization and digit information. These are very useful information to capture the entity extraction. Symbol such as the ',' is found to be very crucial in detecting the boundaries of the NEs. Table 2 lists some of the orthographic features with their examples in uses.
- **Word Normalization:** Two different variants of this feature were developed. The first feature is used to capture the words which are in plural form, verb, alphanumeric letters, hyphen, and digit. This feature actually converts the words to its root form. Another feature represents the orthographic construction of the words. This feature is termed as the word shape feature where each word is mapped to its respective class. Words are normalized by converting every uppercase character by 'A', lowercase character is changed to 'a' and every digit by '0'. For example, if the token is 'Medical-49', its normalized form would be 'Aaaaaaa-00'.
- **Word Class Feature:** This feature is generated motivated by the fact that similar entities reside within the similar class. This feature is the enhancement of the word normalization feature. After the word conversion, the squeezing operation is performed where the consecutive sequence of words is converted into the single character. For example for the word 'Medical-49', its word class form would be 'Aa-0'.
- **Section Feature:** This binary feature was generated after extracting the sections from the training set manually. Top 16 sections were extracted with their probabilities of carrying the relevant PHI terms. These sections were generated manually and form very effective feature in the identification of the PHI terms.
- **Gazetteer List:** For the identification of the PHI terms, task related term list helps in improving the performance of the system. These lists include the headings which give very strong remarks about the occurrence of the PHI terms. Example: 'Transcribed by', 'Dictated by'. These sections carry information about the occurrence of the DOCTOR name. Apart from this, the dictionary was used for the names like the common names across countries, full name and the abbreviation of the common US locations for e.g. "New York" and "NY". We used the list of names for the days, months.

In addition, some lexical clues were identified for each individual PHI category. These are the types of the trigger words that give an indication regarding the presence of the particular PHI terms for example 'Dr.', 'M.D' in the case of DOCTOR, 'Mr.', 'Mrs' in the case of the patient. These lexical clues are obtained from the neighboring contexts of the target PHI terms. These are filtered on the basis of occurrence and their importance with the target PHI categories.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Some of the orthographic features with their examples in uses

Feature	Example	Feature	Example
InitCap	David	CapMixAlpha	EpoR
InCap	mAb	DigitSpecial	12-4
DigitOnly	12,23	AlphaDigitAlpha	IL34S,EIB
DigitAlpha	28BC	CapLowAlpha	Srs, Ras, Epo
Hyphen	_	DigitCommaDigit	2,34
CapsAndDigits	28Houston	LowMixAlpha	mAB,mRNA
StopWord	at, on	AllCaps	DHC
AlphaDigit	s72	EndwithDot	Dr., Mr.

- **General Entity Information:** The Stanford Named Entity Recognizer (http://nlp.stanford.edu/software/CRF-NER.shtml) was used to obtain the named entity tag of the corresponding token.
- **Discourse Features:** This feature plays a very important role in the identification of the terms that are very much related to section headings and their positions of occurrence in the sentence. For example the strong contextual clues words like'CC:' and 'Dictated by'.

Classification: CRF Classifier

We trained our system using the Conditional Random Field algorithm. This algorithm is very popular in the domain of entity recognition because of its ability to tackle the sequence labeling task very efficiently. CRF is categorized under the class of statistical modeling. They are mainly applied for the structural prediction which is formed by the combination of classification and graphical modeling methods. CRF integrates the performance of the graphical models to compactly model multivariate data and classification methods to perform prediction using a large set of input features.

The CRF-classifier works on the Markov Assumption that given the observation sequence of the token x_1, x_2, x_3 .....x_n, the corresponding label-sequence y_1, y_2, y_3 .....y_n is predicted by the CRF. It maximizes the conditional probability for every *y* in the possible set of label sequence. Mathematically, CRF can be represented as

$$\exp\left(\sum_{j} \gamma_j t_j(y_{(i-1)}, y_i, x, i) + \sum_{k} \mu_k s_k(y_i, x, i)\right) \quad (1)$$

Here, $t_j(y_{(i-1)}, y_i, x, i)$ represents the transition feature function for the whole set of observation sequence. *i* and (*i*-1) denote the position of the labels in the label sequence. Here $s_k(y_i,x,i)$ represents the state feature function at the label position *i* and observation sequence. γ_j and μ_k are the parameters estimated from the training dataset. While defining the feature function, the real valued feature set of the observation is constructed as $b(x,i)$. It expresses empirical distribution characteristics of the training data in order to hold the model distribution.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

There are various other algorithms that can handle the sequence labeling task like Maximum Entropy Markov Model (MEMM) and Hidden Markov Model (HMM). The main advantage of CRF over HMM is that it is conditional in nature, which helps in relaxation of the independence assumption as needed by the HMM. Apart from that, CRF is very much efficient in handling the label bias problem. Label bias problem is due to the directed graphical model that is exhibited by MEMM and HMM. It was observed that CRF outperforms the other models in a number of tasks such as speech recognition, bioinformatics.

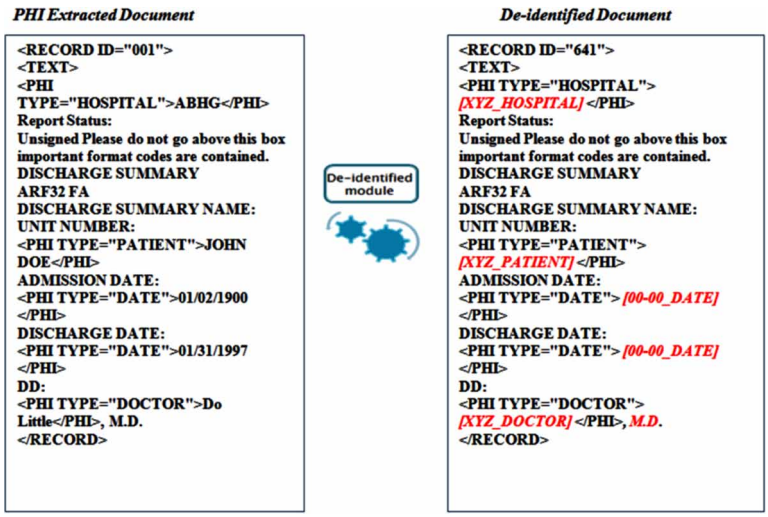
De-identification task targets to automatically extract the PHI terms with the proper boundary identification. As CRF being the better performer than other models, this work used CRF as an underline base classifier. The open source toolkit CRF++ was used to carry out our experiments. BIO notation was followed to label the PHI terms which occur in sequence. Here 'B' identifies the beginning of the PHI token, 'I' represents that the PHI token is within the sequence. 'O' is used to identify that token belongs to none of the PHI term categories.

For the classifier learning, whole feature set discussed in the previous section was used. CRF modeling is done by considering all the PHI categories through learning by the training phase and predicting the output of the model built during the test phase.

De-Identification of PHI Terms

In order to hide the patient personal information, de- identification of the PHI terms is highly necessary. A template was adopted through which we convert all PATIENT, HOSPITAL, DOCTOR names present into a generic format like XYZ_PATIENT, XYZ_HOSPITAL, XYZ_DOCTOR, respectively and all the dates into the format 00-00_DATE. Thus, capturing the information required without compromising the personal details. We represent this below in Figure 3 by considering the original document and the de-identified document. Figure 3 shows De-Identification Process.

Figure 3. De-identification process



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DATASET AND EXPERIMENTS

Dataset

The 2014 i2b2 challenge dataset was used in the current paper, which was taken from the Research Patient Data Repository of Partners' Healthcare. A total of 1304 medical records basically of longitudinal nature were manually annotated. These records were obtained from 297 patients following the annotation guidelines. The dataset was divided into two parts where the first part consists of 790 records of 188 patients that were used as the training dataset and 514 records of 109 patients were test dataset.

i2b2-2014 challenge dataset compromises of 17,045 PHI relevant instances in the training dataset, while the test dataset consists of total 11,462 PHI instances. To ensure the patient confidentiality as much as possible, the challenge aims to identify HIPAA- PHI categories firstly with the added subcategories. They were annotated using seven main PHI categories with the twenty- five associated subcategories. While our experiments cover the seven main PHI categories, i2b2 challenge covers almost all HIPAA defined categories and subcategories. The list of categories, as well as subcategories that were needed to be identified, is as follows:

- NAME (subtypes: PATIENT, DOCTOR,USERNAME)
- PROFESSION
- LOCATION (subtypes: HOSPITAL, DEPARTMENT, ORGANIZATION, ROOM, STREET, CITY, STATE, COUNTRY, ZIP)
- AGE
- DATE
- CONTACT(subtypes: PHONE,FAX,EMAIL,URL,IPADDRESS)
- Ids(SUBTYPES: MEDICAL RECORD NUMBER, HEALTH PLAN NUMBER, SOCIAL SECURITY NUMBER, ACCOUNT NUMBER, VECHILE ID, DEVICE ID, LICENCE NUMBER, BIOMETRIC ID)

In this work, the aim was to identify four different PHI subtypes; PATIENT, DOCTOR, HOSPITAL name and DATE from the above-defined categories. The experiment was performed with only the training dataset.

Evaluation

The system was evaluated by using 10 fold cross-validation on the training data. For evaluating the classifier's performance, we have used the most popular evaluation metrics; recall, precision, and F-Measure. Recall also known as sensitivity is defined as the ratio of total number retrieved relevant documents to the number of relevant records. Precision represents the positive prediction and is the ratio of retrieved relevant records to the total number of documents.

Recall and precision are represented as follows

$$recall = \frac{Total\ number\ of\ correctly\ predicted\ NE}{Total\ number\ of\ actual\ NE\ present\ in\ the\ gold\ data} \tag{2}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

$$precision = \frac{Total\ number\ of\ correctly\ retrieved\ NE}{Total\ number\ of\ NE\ retrieved\ by\ the\ system} \tag{3}$$

F-Measure considers both the precision and recall. It is defined as the harmonic mean of recall and precision. We defined F-Measure value as follows:

$$f_{\alpha} = \frac{(1+\alpha^2)(recall * precision)}{\alpha^2 * precision + recall} \tag{4}$$

Here we set the value of α=1.

Experimental Results

Table 3. reports the extensive results of 10-fold cross validation: Recall, Precision and F-score values for individual entities. The system was developed using the handcrafted feature set trained on CRF. It was an interesting observation that with the set of whole features learned using CRF, the system performs quite well. The system was evaluated using 5,6,7,8,9,10 fold cross validation. Here, the performance of the proposed system with10 fold cross validation is reported.

Table 3 shows that overall system has achieved good recall values as compared to precision values. But there are some interesting observations regarding the recognition of individual entities. If the comparison is done on the performance among 3 important PHI terms; Hospital name, Doctor Name & Patient name then Doctor class is having the highest precision and recall values of 99.70% and 93.80% respectively. Precision and recall values for Date are quite good, but the performance can be further improved by introducing some more date based patterns. The recall value for the hospital is low as compared to all the PHI terms introduced here; the one possible reason behind this is to use abbreviation instead of full hospital names.

The report shows that the system is accurate in predicting the PHI terms with the high degree of precision. The detailed error analysis is performed on the obtained results.

Table 4 shows different error cases. Here, X or Y represents any class like PATIENT, DOCTOR, HOSPITAL, and DATE. “B_X” or “B_Y” represents the beginning of the token, “I_X” or “I_Y” represents the intermediate of the token and “O” denotes others.

Table 3. Results of 10-fold cross validation: Recall, precision & F-score values for individual entities

Class	Recall	Precision	F-Score
Overall	0.9569	0.9931	0.9746
Hospital	0.9056	0.9769	0.9399
Doctor	0.9380	0.9970	0.9666
Patient	0.9355	0.9831	0.9587
Date	0.9905	0.9984	0.9944

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Different error cases

Case	Actual Labels	Classified Labels
MISSED ENTITY	B_X I_X	O
ACTUAL ENTITY	B_X I_X	B_Y I_Y
WRONG BOUNDARY	B_X I_X	B_X O
OVER PREDICT	O	B_X I_X

Table 5 mentions about case 1: missed entity error. Here the system was unable to identify the entity. In the following e.g. 'DO' which in actual represents the Doctor name was predicted as others. The reason for the error is the short words or the abbreviated words. 'DO' is the abbreviated form of doctor name 'DO D LITTLE' which is missed by the system.

Table 6 presents the case 2: actual entity error. Here the entity was predicted but was not able to classify it in the proper class. For example in the following sentence, 'FOR' was captured as the B-DOCTOR while it should be others.

Table 7 represents the case 3: wrong boundary error. This case represents the boundary detection problem, where the system was able to capture the token properly but was not able to properly identify the boundary of the token. In the below example 'LITTLE' was not identified as the boundary of the system.

Table 8 represents the case 4: over predicted error. Last error occurred due to the over prediction where non-entity was categorized as the entity. For example here 'DO LITTLE' was categorized as the entity and was tagged by the DOCTOR. It might be because our system is unable to properly identify the context.

Table 5. CASE 1: Missed entity error

Words	DR.	DO	2	.	DO	D	LITTLE	,	M.D	.
Actual	O	B-DOCTOR	O	O	B-DOCTOR	I-DOCTOR	I-DOCTOR	O	O	O
Predicted	O	O	O	O	B-DOCTOR	I-DOCTOR	I-DOCTOR	O	O	O

Table 67. CASE 2: Actual entity error

Words	Dictating	For	DO	LITTLE	,	M.D	.
Actual	O	O	B-DOCTOR	I-DOCTOR	O	O	O
Predicted	O	B-DOCTOR	I-DOCTOR	I-DOCTOR	O	O	O

Table 7. CASE 3: Wrong boundary error

Words	HE	WILL	FOLLOW-UP	WITH	DR.	DO	LITTLE
Actual	O	O	O	O	O	B-DOCTOR	I-DOCTOR
Predicted	O	O	O	O	O	B-DOCTOR	O

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. CASE 4: Over predict error

Words	DO	LITTLE	THORACOSCOPIC	EXAM	BX	DRAINAGE	.
Actual	O	O	O	O	O	O	O
Predicted	B-DOCTOR	I-DOCTOR	O	O	O	O	O

From the confusion matrix presented in Table 3, it was identified that mainly case 1 type of the error was more dominated.

- DOCTOR was tagged as O where there are the following forms: < D1 > M.D., < D2 > M.D., < D3> M.D.
- Misclassification has occurred when Hospital name is written in short for example CMC, FIH, OLH, HGMC etc.
- Most of the misclassification errors occurred when the patient name was preceded by the Mr., Mrs., Ms.
- Dates of the forms like 11th of April, January 25, Veterans Day were not captured properly.

CONCLUSION

In this paper, a clinical de-identification system is developed that aims to identify the Protected Health Information and to classify them into the predefined categories present in the clinical records. The proposed approach for de-identification is purely based on a popular machine learning technique. A variety of domain independent features, semantic, lexical features were designed for the system to be domain adaptable. Various others features extracted from localization dictionaries were combined in order to further enhance the performance of the system. The proposed system achieves encouraging performance. The system obtained the overall F-Measure of 97.46%, in identifying PATIENT, HOSPITAL, DOCTOR names and DATE.

However, a lot of issues are yet to be resolved. One such issue is to resolve abbreviation problem which is identified as a non-PHI term. In addition, it would be interesting to observe the performance of the proposed system on other type of PHI mentions.

FUTURE WORK

In future, this work can be extended using deep neural network architecture to identify the PHI terms from the medical records. Another major challenge is to deal ambiguity and abbreviation problem to properly distinguish PHI terms with non-PHI terms which can be another future work.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Aramaki, E., Imai, T., Miyo, K., & Ohe, K. (2006, November). Automatic deidentification by using sentence features and label consistency. In *i2b2 Workshop on Challenges in Natural Language Processing for Clinical Data* (pp. 10-11).

Aronson, A. R. (2001). Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In *Proceedings of the AMIA Symposium* (p.17). American Medical Informatics Association.

Beck, T., Gollapudi, S., Brunak, S., Graf, N., Lemke, H. U., Dash, D., ... Brookes, A. J. (2012). Knowledge engineering for health: A new discipline required to bridge the ICT gap between research and healthcare. *Human Mutation*, *33*(5), 797–802. doi:10.1002/humu.22066 PMID:22392843

Beckwith, B. A., Mahaadevan, R., Balis, U. J., & Kuo, F. (2006). Development and evaluation of an open source software tool for deidentification of pathology reports. *BMC Medical Informatics and Decision Making*, *6*(1), 1. doi:10.1186/1472-6947-6-12 PMID:16515714

Berner, E. S., Detmer, D. E., & Simborg, D. (2005). Will the wave finally break? A brief view of the adoption of electronic medical records in the United States. *Journal of the American Medical Informatics Association*, *12*(1), 3–7. doi:10.1197/jamia.M1664 PMID:15492029

BioASQ. (n.d.). *BioASQ project, Data*. Retrieved January 07, 2015, from http://www.bioasq.org/participate/data

Chinchor. (2001, January 12). *Named Entity Scores– English, in: Message Understanding Conference Proceedings.* Retrieved January 06,2015, from http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/ne_english_score_report.html

Coorevits, P., Sundgren, M., Klein, G. O., Bahr, A., Claerhout, B., Daniel, C., ... De. (2013). Electronic health records: New opportunities for clinical research. *The Association for the Publication of the Journal of Internal Medicine*, *274*(6), 547–560. doi:10.1111/joim.12119 PMID:23952476

Dehghan, A., Kovacevic, A., Karystianis, G., Keane, J. A., & Nenadic, G. (2015). Combining knowledge- and data-driven methods for de-identification of clinical narratives. *Journal of Biomedical Informatics*, *58*, S53–S59. doi:10.1016/j.jbi.2015.06.029 PMID:26210359

Deleger, L., Molnar, K., Savova, G., Xia, F., Lingren, T., Li, Q., ... Solti, I. (2013). Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *Journal of the American Medical Informatics Association*, *20*(1), 84–94. doi:10.1136/amiajnl-2012-001012 PMID:22859645

Friedman, C., Shagina, L., Socratous, S. A., & Zeng, X. (1996). A WEB-based version of MedLEE: A medical language extraction and encoding system. In *Proceedings of the AMIA Annual Fall Symposium* (p. 938). American Medical Informatics Association.

Goeuriot, L., Jones, G.J., Kelly, L., Leveling, J., Hanbury, A., Müller, H., Salantera, S., Suominen, H. and Zuccon, G. (2013). ShARe/CLEF eHealth Evaluation Lab 2013, Task 3: Information retrieval to address patients' questions when reading clinical reports. *CLEF 2013 Online Working Notes, 8138*.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hara, K. (2006, November). Applying a SVM based Chunker and a text classifier to the deid challenge. In *i2b2 Workshop on challenges in natural language processing for clinical data* (pp. 10-11).

He, B., Guan, Y., Cheng, J., Cen, K., & Hua, W. (2015). CRFs based de-identification of medical records. *Journal of Biomedical Informatics*, 58. PMID:26315662

Health Insurance Portability and Accountability Act. P.L. 104e191, 42 U.S.C. 1996.

Hirschman, L., Yeh, A., Blaschke, C., & Valencia, A. (2005). Overview of BioCreAtIvE: Critical assessment of information extraction for biology. *BMC Bioinformatics*, *6*(Suppl 1), S1. doi:10.1186/1471-2105-6-S1-S1 PMID:15960821

Kim, J. D., Ohta, T., Pyysalo, S., Kano, Y., & Tsujii, J. I. (2009, June). Overview of BioNLP'09 shared task on event extraction. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing: Shared Task* (pp. 1-9). Association for Computational Linguistics. 10.3115/1572340.1572342

Lafferty, J., McCallum, A., & Pereira, F. C. (2001). *Conditional random fields: Probabilistic models for segmenting and labeling sequence data*. Academic Press.

Liu, H., Lussier, Y. A., & Friedman, C. (2001). Disambiguating ambiguous biomedical terms in biomedical narrative text: An unsupervised method. *Journal of Biomedical Informatics*, *34*(4), 249–261. doi:10.1006/jbin.2001.1023 PMID:11977807

Liu, Z., Chen, Y., Tang, B., Wang, X., Chen, Q., Li, H., ... Zhu, S. (2015). Automatic de-identification of electronic medical records using token-level and character-level conditional random fields. *Journal of Biomedical Informatics*, 58. PMID:26122526

McCallum, A., Freitag, D., & Pereira, F. C. (2000, June). *Maximum Entropy Markov Models for Information Extraction and Segmentation* (Vol. 17). ICML.

Meystre, S. M., Friedlin, F. J., South, B. R., Shen, S., & Samore, M. H. (2010). Automatic de-identification of textual documents in the electronic health record: A review of recent research. *BMC Med Res Methodol BMC Medical Research Methodology*, *10*(1), 70. doi:10.1186/1471-2288-10-70 PMID:20678228

Nédellec, C., Bossy, R., Kim, J. D., Kim, J. J., Ohta, T., Pyysalo, S., & Zweigenbaum, P. (2013, August). Overview of BioNLP shared task 2013. In *Proceedings of the BioNLP Shared Task 2013 Workshop* (pp. 1-7).

Pyysalo, S., Ohta, T., Rak, R., Sullivan, D., Mao, C., Wang, C., ... Ananiadou, S. (2012). Overview of the ID, EPI and REL tasks of BioNLP Shared Task 2011. *BMC Bioinformatics*, *13*(11), 1. PMID:22759456

Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., & Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*, *17*(5), 507–513. doi:10.1136/jamia.2009.001560 PMID:20819853

Stubbs, A., Kotfila, C., & Uzuner, Ö. (2015). Automated systems for the de-identification of longitudinal clinical narratives: Overview of 2014 i2b2/UTHealth shared task Track 1. *Journal of Biomedical Informatics*, *58*, S11–S19. doi:10.1016/j.jbi.2015.06.007 PMID:26225918

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sundheim, B. M. (1996, May). Overview of results of the MUC-6 evaluation. In *Proceedings of a workshop on held at Vienna*, *Virginia* (pp. 423-442). Association for Computational Linguistics. 10.3115/1119018.1119073

Szarvas, G., Farkas, R., & Busa-Fekete, R. (2007). State-of-the-art anonymization of medical records using an iterative machine learning framework. *Journal of the American Medical Informatics Association*, *14*(5), 574–580. doi:10.1197/jamia.M2441 PMID:17823086

Torii, M., Fan, J. W., Yang, W. L., Lee, T., Wiley, M. T., Zisook, D., & Huang, Y. (2014, November). De-identification and risk factor detection in medical records. Seventh i2b2 Shared Task and Workshop: Challenges in Natural Language Processing for Clinical Data.

Uzuner, Ö., Solti, I., & Cadag, E. (2010). Extracting medication information from clinical text. *Journal of the American Medical Informatics Association*, *17*(5), 514–518. doi:10.1136/jamia.2010.003947 PMID:20819854

Uzuner, Ö., South, B. R., Shen, S., & DuVall, S. L. (2011). 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association*, *18*(5), 552–556. doi:10.1136/amiajnl-2011-000203 PMID:21685143

Wellner, B., Huyck, M., Mardis, S., Aberdeen, J., Morgan, A., Peshkin, L., ... Hirschman, L. (2007). Rapidly Retargetable Approaches to De-identification in Medical Records. *Journal of the American Medical Informatics Association*, *14*(5), 564–573. doi:10.1197/jamia.M2435 PMID:17600096

Wolf, M. S., & Bennett, C. L. (2006). Local perspective of the impact of the HIPAA privacy rule on research. *Cancer*, *106*(2), 474–479. doi:10.1002/cncr.21599 PMID:16342254

Xu, H., Stenner, S. P., Doan, S., Johnson, K. B., Waitman, L. R., & Denny, J. C. (2010). MedEx: A medication information extraction system for clinical narratives. *Journal of the American Medical Informatics Association*, *17*(1), 19–24. doi:10.1197/jamia.M3378 PMID:20064797

Xue, H., Chen, S., & Yang, Q. (2008). Structural support vector machine. *Advances in Neural Networks-ISNN*, *2008*, 501–511.

Yang, H., & Garibaldi, J. M. (2015). Automatic detection of protected health information from clinic narratives. *Journal of Biomedical Informatics*, 58. PMID:26231070

KEY TERMS AND DEFINITIONS

Classification: Classification in the machine learning is defined as the supervised learning technique where problem is to identify the class of the new observation with the already developed observations through the labeled data.

Conditional Random Field (CRF): An undirected probabilistic graphic model categorized under statistical modeling method used for structured prediction in machine learning and pattern recognition.

Health Insurance Portability and Accountability Act (HIPAA): An act formulated in 1996 for the secure transmission and usability of the private health information electronically.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Machine Learning: A field of computer science that exploits the development of the algorithms for making a prediction on the data on the basis of its learning.

Medical Natural Language Processing: A domain of natural language processing that focuses on text mining of the medical records like clinical texts.

Named Entity Recognition (NER): An entity extraction task which aims to identify and retrieve the text carrying relevant information into some predefined categories.

Natural Language Processing (NLP): A domain of computer science, computational linguistics and artificial intelligence that focuses on computers understand to human languages or natural language.

Support Vector Machine (SVM): A supervised learning model used in the classification that learns from the data by developing a model that maximizes the error margin.

This research was previously published in the Handbook of Research on Applied Cybernetics and Systems Science edited by Snehanshu Saha, Anand Narasimhamurthy, Abhyuday Mandal, Sarasvathi V, and Shivappa Sangam; pages 234-253; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 76
Mobile Application for Patients' Waiting Time Control and Management of Diagnostic Imaging Examinations

Dimitrios Zarakovitis
University of Peloponnese, Greece

Dimitrios Tsoromokos
University of Peloponnese, Greece

Nikolaos Tsaloukidis
University of Peloponnese, Greece

Athina Lazakidou
University of Peloponnese, Greece

ABSTRACT

Geographic information systems (GIS) are useful informative systems for reducing the waiting time of diagnostic imaging examinations. ArcGIS software is used to develop a digital questionnaire which is used as a data collection tool. The information concerns the patients' personal information, type of examination required and medical history. Data is collected in real time and through GPS capabilities, the patients' exact coordinates are determined. GIS applications are used to create digital maps which display the average waiting time for performing imaging tests. Questionnaire data and interactive maps are uploaded to a digital platform. Through this application, it is possible to locate patients who actually need diagnostic imaging examinations in real time. Observing the location of patients on digital maps makes it easier to redirect them to the nearest health care units with the shortest waiting time. Data can also be acquired by mobile phones.

DOI: 10.4018/978-1-7998-1204-3.ch076

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

In recent years the digitization of health services has greatly optimized the quality of the services provided. The continuous improvement of online applications helps users to deal with the rising health needs and contributes to cost savings. The use of digital media requires the handling and processing of large volumes of data. In the past, most data was stored in hard copy form, but nowadays, big data digital processing is considered to transform the way of living and working in most business areas. By digitizing, potential benefits occur that include early detection of diseases, effective prevention of healthcare fraud and prediction of financial outcomes (Raghupathi & Raghupathi, 2014). Big data analysis helps healthcare providers to develop cost-effective treatments and improve the quality of healthcare services. The release of big data enables health companies to develop innovative applications that help patients to improve their health status. Web-based applications use Informatics Technology (IT) to guide patients to the most effective treatment methods and thorough diagnosis.

Several review papers reveal the advantages of Geographic Information Systems (GIS) in healthcare services. GIS enable users to identify spatial inequalities inaccessibility of healthcare services (Neutens, 2015). They can be used to direct patients in neighboring healthcare services, in a very short time, especially when patients desire specific treatments (Masoodi & Rahimzadeh, 2015). The GIS analysis supports advanced calculations of disease distribution and environmental risk factors (Béland et al., 2014). Furthermore, the majority of review papers report positive results in integrating GIS applications and mobile applications (Nhavoto & Grönlund, 2014). Recent studies also reveal that GIS systems can easily be connected to mobile cloud computing systems for effectivebig data processing (Lo'ai et al., 2016).

The aim of this paper is to identify how mobile technologies and GIS applications can be used for estimating and reducing waiting time for diagnostic tests. The paper proposes a digital application that helps patients to choose the most appropriate hospital unit, in terms of the shortest waiting time, for performing diagnostic imaging examinations, by using their mobile phone. The whole process is done in real time and enables the processing and analysis of dynamic elements. The paper also aims to promote the knowledge and practice in the fields of management of technology in healthcare between health professionals and policy makers.

BACKGROUND

Recent studies focus on the benefits of Informatics Technology (IT) in the public sector. The most commonly used information systems that are used in the public sector are the Geographic Information Systems (GIS). GIS require a great investment in financial and human resources. They are mainly financed by public investors (Kurwakumire, 2014). In most cases, the implementation of GIS technology is used to test the efficiency of healthcare units and hospitals. They are used for analyzing and understanding the spatial dimension of health care (Topol, 2015). GIS tools enable the transformation of quantitative healthcare data to geographic objects. The process is known as geocoding and is used in data analysis (Yasobant et al., 2015). The use of GIS can solve many spatial-orientated management and planning problems (Jawad et al., 2013).

Besides all the significant benefits of using GIS mobile technologies, there are also some limitations. For example, great delays can be encountered when GIS applications access cloud services. There are also many challenges associated with GIS mobile applications and mainly is the privacy protection of

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

owner's data (Mehmood et al., 2015). Some papers reveal more limitations related to GIS processing of sociocultural data. In that case, the constraints refer to data exclusion and data incompatibility (Bagheri, 2014).

The spatial analysis tools enable the evaluation of accessibility in primary healthcare services (Craglia & Maheswaran, 2013). These systems provide sophisticated presentation capabilities, faster big data processing (Amirian et al., 2017), spatial analysis, retrieval and archiving. In several countries, such as Sri Lanka, GIS are used to evaluate the quality of healthcare services in the private sector that mainly provides therapeutic services rather than preventive (Govindaraj et al., 2014). Thus, they can be used to solve problems related to the identification of patients who require immediate diagnostic imaging examinations. They can also aggregate features related to patients' screening examinations and their medical history.

Thus, spatial modeling capabilities (Azevedo, 2016) allow the embedding of digital forms on interactive maps (Júnior et al., 2015). Users can answer key queries by using a geodatabase (Jankowska et al, 2016) that provides high-quality healthcare services. The use of GIS applications includes multidimensional information and concerns patients of different gender, age, level and health needs, as well as access to healthcare services. The spatial modeling of primary information allows the integration of digital questionnaires and interactive maps. Retrieving data in real time is enabled through the web - surveys (Mook et al., 2016) by using the capabilities of GPS and GIS tools. The whole application forms the basis of creating a digital platform for the management and control of the number and type of diagnostic tests. It also optimizes the efficiency of primary health care services using mobile phones, tablets and computers.

A MOBILE APPLICATION FOR PATIENTS' WAITING TIME CONTROL AND MANAGEMENT OF DIAGNOSTIC IMAGING EXAMINATIONS

The main terms that need to be clarified in this study are the following:

1. **Geographic Information Systems (GIS):** They include a database, consisting of observations with spatially distributed elements, processes and events defined in the space as points, lines or surfaces (Longley et al., 2015). They could be described as smart maps that simulate real world (Fradelos et al., 2014). GIS are used for depicting qualitative and quantitative data on maps by using raster or vector files. For example, GIS users can present the total healthcare expenditure of several hospitals on maps;
2. **Diagnostic Imaging Examinations** are the screening examinations that use:
    a. Ionizing radiation such as simple X-rays, mammographs, CT scan, bone densitometry for the diagnosis of osteoporosis (Heidari et al., 2015);
    b. γ-rays, such as bone scan;
    c. Non-ionizing radiation, such as magnetic resonance imaging (MRI) (Busong & Clarke, 2014) and ultrasound imaging (U/S);
    d. A micro camera equipped with a suitable light source used to illustrate the internal organs of the digestive system, such as gastroscopy, colonoscopy, bronchoscopy, etc.;
3. **Geodatabase:** It is defined as a database containing spatial data suitable for processing and analysis (Tsouhlaraki & Achilleos, 2015). A geodatabase is a store of spatial and attributive data. It supports all kinds of attribute data, geographical layers, satellite and aerial imagery (as raster and vector

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

data), land surface 3D modelling data, network systems, and field measurements (Ciampalini et al., 2015). The implementation of a geodatabase includes a collection of files or tables in order to make useful conclusions. For example, users create a geodatabase when they collect data associated with medical equipment. In that case, data is stored in shapefiles spatially distributed on maps. The total of these files is defined as a geodatabase;

4. **Survey 123:** Is a GIS application used for creating digital questionnaires. It is an intuitive form-centric data gathering tool that helps users to create, share and analyze surveys in real time. This application enables users to locate the exact coordinates of the participants in the questionnaire. All answers are saved in a database and can be shared among members of a group by using credentials. Survey 123 is supported by the Environmental Systems Research Institution (ESRI);
5. **Operations Dashboard for ArcGIS:** Is a web application that enables data visualization and analytics for real-time operational view of services and events. It offers analytics for real-time operational view of people, services, assets and events (ESRI, 2017a). For example, it can be used for dynamic monitoring of the activities of medical personnel, the exact position of medical equipment and the number of diagnostic imaging examinations performed in hospitals;
6. **Shapefile (.shp):** Is a spatial data format that can be read by a variety of programs (Löwenberg-Neto, 2014). A shapefile is a simple, nontopological format for storing the geometric location and attribute information of geographic features. Geographic features in a shapefile can be represented by points, lines, or polygons (areas).

METHODS

The proposed application is created in six different steps:

1st Step: Waiting time data for performing mammographs from 12 different Greek public hospitals of the 6th Health Region of Peloponnese was collected. Data were collected by telephone from the radiology departments of the public hospitals. Data were counted in days and was collected during 2017 (Table 1). Waiting time data must be updated monthly, in order to acquire the exact waiting time for each hospital unit. The type of the mammography system was also configured by telephone.

2nd Step: Waiting time data is converted to shapefiles and digitized. That means that all information data of Table 1 is uploaded on GIS platform and converted to symbols on a digital map. By clicking on these symbols, useful windows emerge. These windows contain waiting time data of Greek hospitals. The digital map is updated monthly.

3rd Step: At this point, the mobile application for waiting time control and management of diagnostic imaging examinations is created. The web application consists mainly of an anonymous digital questionnaire. The questionnaire was formed for patients' data collection by using Survey 123for ArcGIS. A variety of questions is digitizedin order to collect accurate patients' data (ESRI, 2017b). In order to create the web application, users choose the desired template and the type of questions (e.g. multiple choice questions) associated with patients. Survey 123enables users to perform a statistical analysis after patients have answered the questions. In that way, they can calculate the number and type of answers. The statistical analysis software supports GIS technologies. In that way, users can determine the exact location of patients who need a specific imaging examination (e.g. a digital mammography). The answers can be stored and shared among users. This application

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

is specially designed by ESRI to run on mobile phones and tablets too. Once the digital questionnaire is created, it is ready for downloading from Google Play Store. Patients download the Survey 123application and register to ESRI website, in order to run the application. By accepting the licenses and terms of the web application, they allow secure personal data processing and sharing of their location. Their location is determined by integrating GPS technology. Patients are asked to complete their gender, age, insurance provider, medical history and the type of the examinations needed. That is a real-time filling up procedure by patients' smartphone.

4thStep: GIS operational dashboard web application was used for directing patients to the nearest hospital units with the shortest waiting times for performing imaging examinations. Dashboards consist of elements, such as charts, maps, gauges, lists, and indicators. A dashboard occupies 100% of the application browser window. Elements can be stacked or grouped in various ways. There are progress and meter gauges that are used to show values moving forward or configure values that show the amount of a variable. The indicators are cards that are intended to count features or summarize the sum or the average of a variable. Dashboards also consist of details elements that show rows or features from a layer and can be toggled on and off. Thus, serial charts are integrated into the dashboard in order to show two or more series of data. That is an advantageous procedure, in case of statistical analysis of participants' answers in the electronic survey (ESRI, 2017b).

5th Step: The mobile data and the digital map created in the 2ndStep are uploaded to the operational dashboard platform. The online platform was installed on a computer unit in the DigiTHEA Lab in the University of Peloponnese. It acts as the central acquisition data system. At this point, users choose the refresh interval time of the operating background which is related to how many seconds are needed to receive the results of the questionnaire from different geographic regions. By modifying the type and color of the symbols, specific areas are blinking on the map. Each blinking symbol corresponds to a new patient's entry with a different type of desired imaging examinations. The central unit is responsible for data processing, analysis, saving and communication with all the participants.

6th Step: It consists of the pilot study of the mobile application. A sample of 12 members of the DigiTHEA Lab is used. Two of the members were located in the central data acquisition system in DigiTHEA Lab. Ten of them pretended to be the patients who requested diagnostic examinations by using their mobile phone. Their age was between 30 - 50 years old. All ten participants were in different places in Peloponnese at the same time. The digital application was tested regarding five different points (Table 2):

- **Speed Data Transmission:** The members of DigiTHEA Laboratory filled in the questionnaire and requested diagnostic examinations of different types. The time for filling in the questionnaire was calculated by using a stopwatch. Requested data was sent by the participants' mobile phone to the central acquisition data system. Most of the participants sent their data almost simultaneously;
- **Clarity of the Questions:** The participants were asked to comment on how clear were the questions (*Clear* corresponds to 100%, *Not Clear* to 0% and *Some Questions Were Clear* corresponds to 50%). The multiple choice questions enable the classification of all the answers in a table;
- **Data Upload:** All participants were asked to upload their medical prescription. The time of data upload was calculated by using a stopwatch;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Accuracy in Locating Patients' Exact Coordinates:** All participants were asked to use the GPS system of their mobile phone (successful tracking corresponds to 100% and failed tracking to 0%;
- **Accuracy in Redirecting Patients to the nearest Hospital Unit with The Shortest Waiting Time:** All participants were asked to comment on the accuracy of the directions given (accurate corresponds to 100%, not accurate corresponds to 0%).

The members who pretended to be patients were able to download and upload photos and pictures on the mobile application. These pictures were uploaded to the central data acquisition system. The participants who were located in the DigiTHEA Lab acquired all available data. They were able to locate the exact origin of all data acquired. That was accomplished because patients had turned the GPS system of their mobile phone on. Members of the DigiTHEA Lab could focus on specific symbols on the map and popup windows were displayed. The windows contained all available data. Data included the type of imaging examination required, average waiting time, contact details and location of the hospital unit, as well as the cost of the examinations. By choosing an appropriate color symbol, a particular cartographic environment was developed that enabled the grouping of quantitative and quality characteristics. The process of visual modelling of spatial data was also accelerated.

Table 1. The waiting data time for performing a different type of mammographies

Hospital	Waiting Time (Days)	Type of Mammography System
General Panarcadian Hospital of Tripolis	200	Digital
General Hospital of Messinia-Hospital Unit of Kalamata	36	Digital
General Hospital of Lakonia-Hospital Unit of Sparta	27	Film
General Hospital of Argolida-Hospital Unit of Argos	4	Film
General Hospital of Korinthos	25	Digitized
General Hospital of Arta	38	Digital
University Hospital of Patra	90	Digital
General Hospital of Ioannina ''Hatzikosta''	120	Digitized
University Hospital of Ioannina	360	Digital
General Hospital of Kefalonia	43	Film
General Hospital – Hospital Unit of Pyrgos	No more Mammographies until the end of 2017	Digitized
General Hospital of Patra – Hospital Unit 'Agios Andreas'		No
General Hospital of Aitoloakarnania – Hospital Unit of Agrinio		No
General Hospital of Preveza		No
General Hospital of Thesprotia – Hospital Unit of Filiates	Closed	Film
General Hospital of Corfu	Closed	Digitized
General Hospital of Lefkada		No
General Hospital of Lixouri		No

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. The results of the five-point test per different categories

Participants	Time for Completing the Questionnaire (sec)	Questions (Clear: 100, not Clear: 0, Somewhere Clear: 20) (%)	Time of Data Upload (sec)	Accuracy in Patients' Location (%)	Accuracy in Redirecting Patients (%)
Participant 1	15	100	40	100	100
Participant 2	20	100	60	100	100
Participant 3	18	100	45	100	100
Participant 4	35	100	65	100	100
Participant 5	42	100	35	100	100
Participant 6	24	100	52	100	100
Participant 7	36	100	50	100	100
Participant 8	58	100	38	100	100
Participant 9	22	100	46	100	100
Participant 10	60	100	44	100	100
Average	33		47,5		

The central acquisition system could locate patients who needed immediate diagnostic imaging examinations. By studying the map with the waiting time data, members of DigiTHEA lab directed patients to the nearest hospital units with the shortest waiting time. The distance and zoom tools facilitated patients' redirection to the nearest hospital units. Members of the DigiTHEA Lab could communicate with patients through the available chat rooms provided by the system. Thus, the patients could give feedback and make their comments on the quality of healthcare services provided.

The designed digital platform functions as a pool of answers that is filled up in real time by the users. By selecting appropriate filters, the central acquisition data system can answer basic queries in a concise time (e.g. display all patients who desire a digital mammography). The system also provides ready queries added to the dynamic thematic background in the form of realistic widgets.

RESULTS

The results of the five points test per category are the following:

- **Speed Data Transmission:** The results show that all participants filled the questionnaire in less than 1 minute. The average time for filling in the questionnaire was 33 sec;
- **Clarity of the Questions:** 100% of the participants agreed that all the questions were entirely clear;
- **Time of Data Upload:** All participants uploaded their medical prescription in almost a few seconds. The central acquisition data system also achieved fast data processing in the DigiTHEA Lab. The average upload time was 47,5 sec. It is evident that the time of data upload was longer than the time of data transmission;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Accuracy in Locating Patients:** 100% of the participants successfully used the GPS system of their mobile phone. The central acquisition data system successfully managed to locate the participants' exact coordinates with no technical problems. The recording of all acquired patients' data was also successful;
- **Accuracy in Redirecting Patients:** The central acquisition data system provided the participants with exact orders, in order to redirect them to the nearest hospital units with the shortest waiting time.

The pilot study also reveals that:

- The radiology department in the University Hospital of Ioannina, the General Panarkadian Hospital of Tripolis and the General Hospital of Ioannina ('Hatzikosta') have the longest waiting time for performing mammographs (360, 200, 120 days accordingly). In the University Hospital of Ioannina, the average waiting time is almost one year which is the longest time in the 6th Health Region in Peloponnese. It is well mentioned that in case of emergency, in the University Hospital of Ioannina ('Hatzikosta'), the waiting time is reduced in 2 - 5 days;
- The radiology department in the General Hospital of Argolida - Hospital Unit of Argos has the shortest waiting time for performing mammographs (almost four days). It is well mentioned that this hospital unit has a filming mammography system;
- The average waiting time for performing mammographs in the majority of hospitals is almost one month;
- The majority of the radiology departments of the public hospitals in Peloponnese has a digital or digitized mammography system;
- The radiology department of the General Hospital of Argolida - Hospital Unit of Argos, the General Hospital of Laconia - Hospital Unit of Sparta, the General Hospital of Thesprotia - Hospital Unit of Filiates and the General Hospital of Kefalonia have a filming mammography system.

It is well mentioned that 5 Greek public hospitals in the 6th Health Region of Peloponnese do not have a mammography system in their radiology departments. These hospitals are the following:

- General Hospital of Patra - Hospital Unit 'Agios Andreas';
- General Hospital of Aitoloakarnania - Hospital Unit of Agrinio;
- General Hospital of Preveza;
- General Hospital of Lefkada; and
- General Hospital of Lixouri.

Patients who live in these regions are directed to the nearest hospital units equipped with a mammography system. The results of the pilot study show that 2 Greek public hospitals have a mammography system, but they do not perform mammographs because of a lack of personnel. These hospitals are the following:

- General Hospital of Corfu; and
- General Hospital of Thesprotia - Hospital Unit of Filiates.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Furthermore, the General Hospital of Ilia - Hospital Unit of Pyrgos will start arranging new appointments for mammographs at the beginning of 2018. The main reason for not closing any more appointments was the high demand of diagnostic imaging examinations added to the lack of personnel.

DISCUSSION

Health policy makers face a compelling need to assess the comparative value between the cost and the income of diagnostic imaging examinations (Neumann et al., 2014). This issue is added to the increased need for reduced cost and shortest waiting time of screening tests. Patients' satisfaction represents their attitude towards care and is strongly associated with the waiting time for performing screening tests (Maseko et al., 2014).

Several studies focus on creating a measurement framework for reducing the waiting time of treatment and diagnosis (Gilbert et al., 2014). Thus, some studies emphasize on creating mobile applications used mainly for reducing the time of diagnosis (Laksanasopin et al., 2015). Some researchers argue that mobile applications are not as helpful as expected and many patients express negative views about them (Musiat et al., 2014). Although this is may be the case, most of the studies do not take into account the waiting time of performing diagnostic exams. They usually focus on reducing the treatment time.

The proposed application forms the basis for reducing the waiting time of screening tests. It promotes the rational management of the number and type of examinations. By using a mobile phone, a tablet or a computer, the patients can inform the central acquisition data system about the examinations needed (Wu et al., 2014). That is a real-time filling up procedure by patients' smartphone (Vij & Shankari, 2015). Patients' requests are forwarded to this center that suitably selects the hospital unit to be addressed to. In that way, the decongestion of healthcare units, namely of the laboratory hospital departments, is enhanced and the process of closing the appointments is accelerated (Chiuchisan et al., 2014). Health managers can easily record the exact number of all the diagnostic examinations performed in public hospitals and display the results on interactive maps. The results can be related to the total expenses, and income of the imaging examinations and the outcomes can be presented on 'smart' maps by using GIS tools (Huang, 2017). The integration of charts and report tables improves the cost-effective analysis (De La Torre - Díez, 2015).

The completion of the digital questionnaire can be done very quickly. Patients do not have to spend much time on arranging their appointments in the secretarial support departments. Within a short period, users can be informed about the patients' medical history and inform the diagnostic imaging departments accordingly. Patients' feedback on the proposed application will improve the quality of health care services. The mobile application will establish channels of communication between health policy - makers, health professionals and patients (Mumtaz et al., 2014).

Thus, the mobile application aims to reduce the waiting time of performing imaging examinations. Greek public hospitals do not provide any official information about the waiting time of performing a diagnostic examination (Economou et al., 2014). The telephone communication ensures the required waiting time data collection. It is clear that Greek public hospitals have not adopted an informatics system for waiting time control until now. Greek patients are discouraged from visiting the radiology departments of public hospitals for performing their diagnostic tests because of the long waiting times (Economou, et al., 2014). Doctors usually advise their patients to perform a digital mammography because of its low radiation and better sensitivity in early breast detection (Smith, 2015). That is the main reason why

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Greek public hospitals with a filming mammography system have shorter waiting times for performing a breast imaging test (e.g. the Hospital Unit of Argos). The second reason for the long waiting time is a lack of health professionals in most Greek Hospitals (Simou et al, 2015).

The economic austerity has caused many problems in the health sector. Recent surveys show that Greek patients who have cancer seem to be satisfied from the healthcare services provided in hospitals (Pini et al., 2014). In most cases, patients' satisfaction is related to the excellent behavior of the health professionals despite the lack of personnel. OECD (Organization for Economic Cooperation and Development) health statistics show that the majority of women in OECD countries have performed a breast cancer screening in 2012 - 2016. Data concern the number of women, 50 - 69 aged, who have performed bilateral mammography divided by the number of women of the same age, answering survey questions on mammography. Health data show that in Sweden in 2014, 90,4% of females, 50 - 69 aged, has performed a breast cancer screening. That is the high est percentage rate among all OECD countries in 2014. In Greece, in the same year, 59,6% of females, between 50 - 69 years old, have performed mammographs. It is clear that there is a growing demand for diagnostic imaging examinations in the majority of OECD countries. Thus, it is well mentioned that a small number of women between 50 - 69 years old in Latvia, Slovak Republic and Turkey have performed mammographs between 2012 -2016. In 2016, the available data show that the number of women in New Zealand, Slovenia and the United Kingdom who have performed mammographs is over 70% (71,5%, 76,8% and 75,4% accordingly).

As far as the Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) exams are concerned, Greece is among the OECD countries with the high estimated number of total CT and MRI exams per 1.000 inhabitants in 2012. In Greece, the average number of total CT exams is 181 per 1.000 inhabitants compared to 256,7 CT exams in the United States in the same year. The number of complete MRI exams per 1.000 inhabitants, in Greece, is 67,9% compared to the total number of MRI exams per 1.000 inhabitants, in the United States (104,8 in 2012).

Furthermore, in Greece, there hasn't been any official waiting list for surgery procedures yet. In most cases, the waiting time of performing mammographs in Greece is longer or almost equal to the waiting list of undertaking a hip replacement in the majority of OECD countries. The waiting list for undertaking a surgery procedure concern the waiting time for specialist assessment to treatment calculated in days. For example, the average waiting time of a hip replacement in Estonia is almost one year in 2016 that is equal to the average waiting time of performing a mammography in the University Hospital of Ioannina. The average waiting time of a hip replacement in Hungary, in 2016, is almost 120 days that is equal to the average waiting time of performing mammographs in the General Hospital of Ioannina ('Hatzikosta'). There is also a growing rise in the waiting time of a hip replacement in Poland in 2012 - 2016. The average increase is almost 20,5% in the same period (OECD, 2016).

CONCLUSION

The primary purpose of the proposed mobile application is to help patients perform diagnostic examinations in a concise time by choosing the nearest hospital unit. It is aimed to help policymakers adopt new guidelines for shortening the waiting times in the laboratory departments of public hospitals. The implementation of digital applications is a practical solution for the decongestion of public hospitals. By using

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

a mobile phone or a tablet, the access to primary health care services is enabled, and a communication channel between health professionals and policymakersis enabled. Thus, it will help health policymakers to create a digital file that contains the waiting time data of the different type of examinations.

The mobile application provides pervasive and timely access to patients' data for all medical professionals. They can retrieve the information needed, such as patients' medical history or pharmaceutical treatment, from the central data acquisition system in real time. They can also perform statistical analysis of the results of screening tests by using GIS tools. The web application enables medical staff to check patients' diagnostic exams in a concise time. Data acquired by the digital questionnaire are available for medical professionals and can be stored in digital files. In that way, the diagnosis and treatment will be accelerated, and patients' medical history can be secured. Medical professionals can organize their tasks more effectively and intuitively. Their working practice will also be improved because they will be able to organize their tasks more effectively and intuitively. The proposed application enables patients to be aware of the waiting time data in advance. Thus, less complaint will be made about the compliance of appointment.

The proposed mobile application combined with the operational dashboard platform is innovative because it allows real-time location of patients who want to perform diagnostic examinations. Existing systems do not allow the geographical patients' location, in addition, to a real-time monitoring of waiting time data. Existing mobile applications take into account the exact patients' coordinates, but they cannot provide dynamic monitoring of waiting time data. Furthermore, the proposed mobile application enables real-time statistical analysis of all the examinations needed in a territory. That is an ideal way of rational management and surveillance of all the examinations performed in hospital units.

The usability of the system ensures the simultaneous rational control of the number, cost and waiting time of imaging examinations. The patients are directed to the hospital units with the shortest waiting time for performing the desired exams. The proposed application provides fast spatial analysis and processing of a large volume of data by using a mobile phone. This kind of data is associated with patients' medical history, images of diagnostic exams (e.g. patients can upload a digital file of an X-ray), the type of examination needed, the cost of examinations and the waiting time data.

Through the proposed application, the assessment of healthcare services will be ensured in order to provide a healthcare system of high quality. The digital platform will provide the basis for the decongestion of hospital units from patients who perform imaging tests by classifying them in urgent and non-urgent cases. That will help saving costs and optimizing the efficiency of healthcare units. It is quite evident that Greek patients have to wait for a long time before performing a diagnostic imaging examination in public hospitals. The total income and expense of diagnostic imaging examinations can be recorded and saved in a database. The whole process is displayed on interactive maps that are easily accessed by health professionals.

The mobile application could be established in the Ministry of Health, and especially in the Greek Health Regions or any hospital units for pilot study. Although there are some cost limitations, ESRI offers plenty of open data and software that can be used to set this application with no charge.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Amirian, P., Lang, T., & Van Loggerenberg, F. (Eds.). (2017). *Big Data in Healthcare: Extracting Knowledge from Point-of-Care Machines*. Springer. doi:10.1007/978-3-319-62990-2

Azevedo, M. J. (Ed.). (2015). *The State of Health and Health Care in Mississippi*. Univ. Press of Mississippi.

Bagheri, N. (2014). What qualitative GIS maps tell and don't tell: Insights from mapping women in Tehran's public spaces. *Journal of Cultural Geography*, *31*(2), 166–178. Retrieved 06 March 2018 from http://www.tandfonline.com/doi/abs/10.1080/08873631.2014.906848 doi:10.1080/08873631.2014.906848

Béland, D., Blomqvist, P., Andersen, J. G., Palme, J., & Waddan, A. (2014). The universal decline of universality? Social policy change in Canada, Denmark, Sweden and the UK. *Social Policy and Administration*, *48*(7), 739–756. Retrieved 06 March 2018 from http://onlinelibrary.wiley.com/doi/10.1111/spol.12064/full doi:10.1111pol.12064

Bushong, S. C., & Clarke, G. (2014). *Magnetic resonance imaging: physical and biological principles*. Elsevier Health Sciences.

Chiuchisan, I., Costin, H. N., & Geman, O. (2014, October). Adopting the internet of things technologies in health care systems. In *2014 International Conference and Exposition on Electrical and Power Engineering (EPE)* (pp. 532-535). IEEE. 10.1109/ICEPE.2014.6969965

Ciampalini, A., Raspini, F., Bianchini, S., Frodella, W., Bardi, F., Lagomarsino, D., ... Onori, R. (2015). Remote sensing as tool for development of landslide databases: The case of the Messina Province (Italy) geodatabase. *Geomorphology*, *249*, 103–118. doi:10.1016/j.geomorph.2015.01.029

Craglia, M., & Maheswaran, R. (Eds.). (2016). *GIS in public health practice*. CRC press.

De La Torre-Díez, I., López-Coronado, M., Vaca, C., Aguado, J. S., & de Castro, C. (2015). Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: A systematic review. *Telemedicine Journal and e-Health*, *21*(2), 81–85. Retrieved 02 February 2018 from http://online.liebertpub.com/doi/abs/10.1089/tmj.2014.0053 doi:10.1089/tmj.2014.0053 PMID:25474190

Economou, C., Kaitelidou, D., Katsikas, D., Siskou, O., & Zafiropoulou, M. (2014). Impacts of the economic crisis on access to healthcare services in Greece with a focus on the vulnerable groups of the population1. *Social Cohesion and Development, 9*(2), 99-115. Retrieved 02 February, 2018, from http://ejournals.epublishing.ekt.gr/index.php/SCAD/article/view/8880

Economou, C., Kaitelidou, D., Kentikelenis, A., Sissouras, A., & Maresso, A. (2014). The impact of the financial crisis on the health system and health in Greece. *Economic crisis, health systems and health in Europe: country experience. Copenhagen: WHO/European Observatory on Health Systems and Policies*. Retrieved 05 February, 2018, from http://www.kentikelenis.net/uploads/3/1/8/9/31894609/economou2014-the_impact_of_the_financial_crisis_on_health_and_the_health_system_in_greece.pdf

Environmental Systems Research Institute. (2017a). *Operations Dashboard*. Retrieved 06 December, 2017, from http://www.esri.com/products/dashboard

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Environmental Systems Research Institute. (2017b). *SURVEY 123for ArcGIS*. Retrieved 06 December, 2017, from http://www.esri.com/products/survey123

Fradelos, E., Papathanasiou, I., Mitsi, D., Tsaras, K., Kleisaris, C., & Kourkouta, L. (2014). Health Based Geographic Information Systems (GIS) and their Applications. *Journal of Academy of Medical Sciences of Bosnia and Herzegovina*, *22*(6), 402–405. Retrieved 14 September 2017. doi:10.5455/aim.2014.22.402-405 PMID:25684850

Gilbert, J. E., Dobrow, M. J., Kaan, M., Dobranowski, J., Srigley, J. R., Jusko Friedman, A., & Irish, J. C. (2014). Creation of a diagnostic wait times measurement framework based on evidence and consensus. *Journal of Oncology Practice / American Society of Clinical Oncology*, *10*(5), e373–e379. Retrieved 04 March 2018 from http://ascopubs.org/doi/abs/10.1200/JOP.2013.001320 doi:10.1200/JOP.2013.001320 PMID:25074879

Govindaraj, R., Navaratne, K., Cavagnero, E., & Seshadri, S. R. (2014). HEALTH CARE IN SRI LANKA: What Can the Private Health Sector Offer? Retrieved 02 March 2018 from http://documents.worldbank.org/curated/en/423511468307190661/pdf/899540WP0Box380th0Care0in0Sri0Lanka.pdf

Heidari, B., Hosseini, R., Javadian, Y., Bijani, A., Sateri, M. H., & Nouroddini, H. G. (2015). Factors affecting bone mineral density in postmenopausal women. *Archives of Osteoporosis*, *10*(1), 15. doi:10.100711657-015-0217-4 PMID:25972061

Hernandez, D. (2014). Artificial intelligence is now telling doctors how to treat you. *Kaiser Health News*. Retrieved 06 December, 2017, from www.wired.com/2014/06/ai-healthcare/

Huang, F., & Gan', L. (2017). The impacts of China's urban employee basic medical insurance on healthcare expenditures and health outcomes. *Health Economics*, *26*(2), 149–163. Retrieved 06 March 2018 from http://onlinelibrary.wiley.com/doi/10.1002/hec.3281/full doi:10.1002/hec.3281 PMID:26524988

Jankowska, M., Schipperijn, J., Kerr, J., & Altintas, I. (2016). Applied CyberGIS in the Age of Complex Spatial Health Data, in *International Conference on GIScience Short Paper Proceedings,* Vol. 1, No. 1. Retrieved 06 March 2018 from https://cloudfront.escholarship.org/dist/prd/content/qt1jr9168v/qt1jr9168v.pdf

Jawad, I. F., Shabana, T., & El-Bakry, M. (2013). Reducing Waiting Time for Transportation using GIS. *International Journal of Advanced Research in Computer Science & Technology*, *4*(1), 62–71.

Júnior, W.J.R., Roig, H.L., Koutrakis, P. (2015). A novel land use approach for assessment of human health: the relationship between urban structure types and cardiorespiratory disease risk. *Environment international*, *85*, 334-342. Retrieved 05 March, 2018, from doi:10.1016/j.envint.2015.09.026

Kurwakumire, E. (2014). Towards a Public Sector GIS Evaluation Methodology. *South African Journal of Geomatics, 3*(1), 33-52. Retrieved 01 March, 2018, from www.ajol.info/index.php/sajg/article/view/106128

Laksanasopin, T., Guo, T. W., Nayak, S., Sridhara, A. A., Xie, S., Olowookere, O. O., . . . Chin, C. D. (2015). A smartphone dongle for diagnosis of infectious diseases at the point of care. *Science translational medicine, 7*(273). Retrieved 01 March, 2018, from stm.sciencemag.org/content/7/273/273re1.short

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lo'ai, A. T., Mehmood, R., Benkhlifa, E., & Song, H. (2016). Mobile cloud computing model and big data analysis for healthcare applications. *IEEE Access, 4*, 6171-6180. Retrieved 02 March, 2018, from ieeexplore.ieee.org/abstract/document/7576619/

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic information science and systems*. John Wiley & Sons.

Löwenberg-Neto, P. (2014). Neotropical region: a shapefile of Morrone's biogeographical regionalisation. *Zootaxa, 3802*(2), 300. Retrieved 04 March, 2018, from www.researchgate.net/profile/Peter_Loewenberg-Neto/publication/262696407_Neotropical_region_A_shapefile_of_Morrone's_2014_biogeographical_regionalisation/links/0c96053c016d0b1dde000000/Neotropical-region-A-shapefile-of-Morrones-2014-biogeographical-regionalisation.pdf

Maseko, F. C., Chirwa, M. L., & Muula, A. S. (2014). Client satisfaction with cervical cancer screening in Malawi. *BMC health services research, 14*(1), 420. Retrieved 05 March, 2018, from bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-14-420

Masoodi, M., & Rahimzadeh, M. (2015). Measuring access to urban health services using Geographical Information System (GIS): a case study of health service management in Bandar Abbas, Iran. *International journal of health policy and management, 4*(7), 439. Retrieved 01 March, 2018, fromhttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4493584/

Mehmood, R., Faisal, M. A., & Altowaijri, S. (2015). Future networked healthcare systems: a review and case study. In *Handbook of Research on Redesigning the Future of Internet Architectures* (pp. 531-558). Hershey, PA: IGI Global. Retrieved 01 March, 2018, from www.igi-global.com/chapter/future-networked-healthcare-systems/131381

Mook, P., Kanagarajah, S., Maguire, H., Adak, G. K., Dabrera, G., Waldram, A., ... Oliver, I. (2016). Selection of population controls for a Salmonella case-control study in the UK using a market research panel and web-survey provides time and resource savings. *Epidemiology and Infection*, *144*(6), 1220–1230. doi:10.1017/S0950268815002290 PMID:26493476

Mumtaz, S., Huq, K. M. S., & Rodriguez, J. (2014). Direct mobile-to-mobile communication: Paradigm for 5G. *IEEE Wireless Communications, 21*(5), 14-23. Retrieved 05 March, 2018, from ieeexplore.ieee.org/abstract/document/6940429/

Musiat, P., Goldstone, P., & Tarrier, N. (2014). Understanding the acceptability of e-mental health-attitudes and expectations towards computerised self-help treatments for mental health problems. *BMC psychiatry, 14*(1), 109 Retrieved 05 March, 2018, from bmcpsychiatry.biomedcentral.com/articles/10.1186/1471-244X-14-109

Neumann, P. J., Cohen, J. T., & Weinstein, M. C. (2014). Updating cost-effectiveness-the curious resilience of the $50,000-per-QALY threshold. *New England Journal of Medicine, 371*(9), 796-797.Retrieved 06 March, 2018, from ieeexplore.ieee.org/abstract/document/6940429/

Neutens, T. (2015). Accessibility, equity and health care: review and research directions for transport geographers. *Journal of Transport Geography, 43*, 14-27. Retrieved 06 March, 2018, from www.sciencedirect.com/science/article/pii/S0966692315000034

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nhavoto, J. A., & Grönlund, Å. (2014). Mobile technologies and geographic information systems to improve health care systems: a literature review. *JMIR mHealth and uHealth, 2*(2). Retrieved 06 March, 2018, from www.ncbi.nlm.nih.gov/pmc/articles/PMC4114429/

Organization for Economic Co-operation and Development. (2016). *OECD. Stat 2016*. Retrieved 10 December 2017 from http://stats.oecd.org//Index.aspx?QueryId=49344#

Pini, A., Sarafis, P., Malliarou, M., Tsounis, A., Igoumenidis, M., Bamidis, P., & Niakas, D. (2014). Assessment of patient satisfaction of the quality of health care provided by outpatient services of an oncology hospital. *Global Journal of Health Science*, *6*(5), 196. Retrieved 01March 2018. doi:10.5539/gjhs.v6n5p196 PMID:25169006

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems, 2*(1), 3. Retrieved 06 March, 2018, fromhttps://www.biomedcentral.com/track/pdf/10.1186/2047-2501-2-3?site=hissjournal.biomedcentral.com

Simou, E., Karamagioli, E., & Roumeliotou, A. (2015). Reinventing primary health care in the Greece of austerity: The role of health-care workers. *Primary Health Care Research and Development*, *16*(1), 5–13. doi:10.1017/S1463423613000431 PMID:24252634

Smith, A. P. (2015). Fundamentals of digital mammography. In *Physics, technology and practical considerations.* Retrieved 06 December, 2017, from www.hologic.com/oem/pdf

Tsouhlaraki, A., & Achilleos, G. (2015). *Learning GIS in practice ArcGIS 9.3*. Athens: Disigma.

Vij, A., & Shankari, K. (2015). When is big data big enough? Implications of using GPS-based surveys for travel demand analysis. *Transportation Research Part C, Emerging Technologies*, *56*, 446–462. doi:10.1016/j.trc.2015.04.025

Wu, H. C., Chang, C. J., Lin, C. C., Tsai, M. C., Chang, C. C., & Tseng, M. H. (2014). Developing screening services for colorectal cancer on Android smartphones. *Telemedicine Journal and e-Health*, *20*(8), 687–695. Retrieved 06 March 2018 from http://online.liebertpub.com/doi/abs/10.1089/tmj.2013.0288 doi:10.1089/tmj.2013.0288 PMID:24848873

Yasobant, S., Vora, K.S., Hughes, C., Upadhyay, A., Mavalankar, D.V. (2015). Geovisualization: a newer GIS technology for implementation research in health. *Journal of Geographic Information System, 7*(1), 20. doi:10.4236/jgis.2015.71002

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 7(4); edited by Anastasius Moumtzoglou; pages 20-33; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 77
Ensuring Privacy of Participants Recruited via Social Media: An Australian Retrospective Visualisation and Roadmap

Chandana Unnithan
https://orcid.org/0000-0002-1832-8499
Torrens University Australia, Australia

Paula M Swatman
Swinburne University, Australia

Jo-Anne Kelder
https://orcid.org/0000-0002-8618-0537
University of Tasmania, Tasmania

ABSTRACT

Researchers worldwide are increasingly looking to recruit research participants via social media (particularly @Facebook and @Twitter) because they appear to offer access to a wider range of research participants and afford inherently convenient tools for recruitment. In Australia, the National Statement on Ethical Conduct in Human Research, together with the federal Privacy law and a number of state-based privacy statutes, provide support and guidance for this novel approach. This article offers a preliminary analysis and discussion of this trend from an Australian perspective, illustrated by an enquiry into the ethical challenges posed by social media-based recruitment, conducted in an Australian university in 2015. Leximancer™ was used as an analytical tool and the content from social media sites used for a small number of research studies conducted up to 2015, taken in conjunction with the various national human research ethics guidelines, offered a means of understanding how ethical challenges of privacy and anonymity can be addressed for responsible social media-based research.

DOI: 10.4018/978-1-7998-1204-3.ch077

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The fundamental principle that people who take part in research projects should be willing and informed participants was articulated as a response to (predominantly medical and psychological) experiments undertaken during the 20th century where participants were anything but willing and informed. Human Research Ethics was designed to protect participants involved in medical research (see, for example, UNLV, 2016). The first attempt to codify the important issues for human participants in medical research was the Nuremberg Code of 1947 (BMJ, 1966a) which emerged from the Doctors' Trials immediately post-WWII. Subsequently, two additional sets of principles – the 1964 World Medical Association's Declaration of Helsinki (BMJ,1966b) and the 1979 U.S. Belmont Report (HHS.gov, 1979) – further clarified and extended the fundamentals of protecting human subjects of research.

Several countries, including the USA, UK, Australia, New Zealand and Canada have since developed their own national set of guidelines for ethical research. Over time, these have expanded from their original, somewhat narrow, focus on medical research to cover social science research; and continue to expand still further to include other fields.

The Australian approach to ethical oversight of research involving humans is centred around Human Research Ethics Committees (HRECs) (NHMRC, 2015). HRECs review research proposals involving human participants to ensure they are ethically acceptable and accord with relevant standards and guidelines. In undertaking this role, Australia's 250+ HRECs strive to ensure the projects they review comply with relevant standards, most importantly the *National Statement on Ethical Conduct in Human Research* (2007, most recently updated in 2018 via a system of rolling reviews). The *National Statement* informs all Human Research Ethics (HRE) in Australia and provides the fundamental guidelines with which all research ethics applications must comply – whether submitted by a university, a hospital, a government agency, a local council, or a not-for-profit organization (NHMRC, 2015).

Australian HRE is governed by both the national statement as well as by the NHMRC's research integrity guidelines, the *Australian Code for the Responsible Conduct of Research* (NHMRC, 2007), together with the federal Privacy Act 1988 and any relevant State privacy acts; plus additional guidelines on specialist issues, such as research involving Aboriginal and Torres Strait Islander (ATSI) participants or issues (Values and Ethics: Guidelines for Ethical Conduct in ATSI Health Research (NHMRC 2003) and Guidelines for Ethical Research in Australian Indigenous Studies (AIATSIS, revised 2/E 2012). Some overlap between research ethics and research integrity is unavoidable and a number of the research integrity requirements identified by the Australian Code (most notably management of research data, peer review and conflict of interest) are also covered by the National Statement and, consequently, incorporated into HRE application forms.

The national statement was originally published in 1999 by the National Health and Medical Research Council (NHMRC, 2015) and, following significant revisions to the entire document in 2007 and 2013, is now updated on a section by section basis to reflect current thinking. For example, Section 3 relating to research methods was reviewed during 2016-17, with an updated version of this Section published in mid-2018: "the draft chapters provide advice for both researchers and … HRECs addressing ethical considerations in the design, development, review and conduct of research (Chapter 3.1*)* [which now incorporates much of the material previously associated only with clinical trials] and ethical considerations specific to genomic research *(*Chapter 3.5*)*" (NHMRC, 2016).

One of the most significant changes in approach by the National Statement has been the expansion of its coverage. Originally focusing predominantly on bio-medical research, the National Statement

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

now also covers research in the social sciences, the arts and humanities; and more technical disciplines such as engineering and Information Technology. In addition to the complexities of expanded research domains, HRE is also facing a second type of complexity as the means of communicating with research participants moves inexorably online. While the addition of email as a communication medium was relatively easy to incorporate into existing HRE guidelines, the emergence of social media is proving far more challenging. The guiding principles of ethical conduct in research – academic merit and integrity; justice; beneficence; and respect (NHMRC, 2007) – are relatively well-understood by those reviewing 'traditional' research projects, but not nearly so straightforward for HRECs grappling with online research approaches and projects. A more nuanced understanding of how these principles apply in new and complex technology-mediated social spaces is required. Henderson, Johnson and Auld (2013) argue that when social media is involved, available guidelines for moral deliberations can be unclear – for example: who are the participants? who owns the data? what is private? and what are the consequences now and in the future?

This paper is organised as follows. Section 2 introduces the background to research ethics and social media (i.e. the historical context), recruitment via social networks, the notion of privacy, and related concerns regarding security as well as ownership of data. Section 3 details the research design and methodology used for the exemplary study. Methods of data collection, coding and analysis are presented. The section also introduces the text analytic tool used namely, Leximancer™. Section 4 presents the visualisation of data analysis using Leximancer™ and the discussion of emerging themes, drawing into a summary. Section 5 presents the conclusions from this study, together with a possible future outlook on social media's challenges to HRE, as well as recommendations for future research on this topic.

RESEARCH ETHICS AND SOCIAL MEDIA RECRUITMENT

The evolution of Information Communication Technologies and the Internet has had a significant impact on the recruitment of participants for research studies. The Internet supports research via two distinctly different approaches: as a tool for finding and accessing people and material (using search engines or online surveys); and as a location for meeting and interacting with people and groups. While these differing online uses are quite self-evident for the 'traditional' Internet, this dual usage is by no means intuitively obvious when considering social media.

Users of social media do not fit neatly into categories, as tools are increasingly embedded into social networking platforms (such as @Facebook or @Twitter), blurring the boundaries between recruiting activities and data collection activities over time (Buchanan, 2012). For example, @Facebook may be used as a *tool* for searching a person's profile, asking them to undertake an embedded survey or for connecting the person with others/groups of people existing on it.

Meeker (2017) in the latest Internet Trends Report, for example, states that @Facebook and @Google/Alphabet (as @Twitter is owned by @Google) now control 85% of the growth in online advertising and, as a consequence, the boundaries between Buchanan's formerly intuitive concepts of tool vs. location are increasingly difficult to identify for research using social media.

Nonetheless, researchers in Australia (and in other countries) cannot ignore the potential social media tools offer for research recruitment. According to Sensis (2017) the most recent survey of social media use in Australia indicates that:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- 79% of Australian Internet users are using social media sites, up from 69% in 2016;
- 59% of users access social media every day (or most days) and 35% check their social media accounts more than 5 times each day;
- 81% of social media users now access their accounts via smartphone (up from 72% a year previously) and 67% of this group access the Internet 5+ times a day
- 94% use @Facebook, 46% use @Instagram, 18% use @LinkedIn, 40% use Snapchat, 32% use @Twitter and 51% use YouTube;
- The average use spends 10 hours each week on @Facebook – a decline of around 2.5 hours per week since 2016 – but more time is now being spent on @Twitter, @Instagram and @LinkedIn;
- The percentage of users following businesses or brands via social media has declined from 1 in 3 to 1 in 4 over the past year across all age segments (except 30-39)
- Attitudes towards social media advertising have become more positive and likelihood to trust brands interacting positively with customers on social media has increased from52% to 64%

Recruitment via Social Networks

Researchers venturing into social media for research purposes need to pay meticulous attention to recruitment of participants (Buchanan, 2012). The crucial question in a recruitment medium where identity is not always clear may well be: Do you really know who your participants are? Is your apparently 18-year-old female participant actually a 55-year-old male? Or even a group of 12-year-old children sharing a single computer? How can you ensure participant privacy, confidentiality and anonymity in an online environment where 'sharing' is so popular and so widely used? Research participants must give informed consent, which is extremely difficult to ensure on social media. Buchanan (2012) argued that informed consent can be obtained via age verification and documentation, but other writers (see, for example, Miller-Merrell, 2012 or Gelinas, Pierce et al., 2017) note that such identification is exceedingly hard to guarantee.

A further concern is the representativeness of social media: Can social media represent the topic/issue under investigation? This really depends on whether the specific social medium being used for recruitment or study is typical of the user group in question. @Twitter is a particularly good example, with both academic and business commentators noting the platform's lack of representativeness (Mitchell & Guskin, 2013; Pack, 2015; Blank, 2016).

Finally, we need to consider issues of data sharing, storage, terms of service and end-user license agreements. When recruiting participants for a research project via social media the nature of the medium ensures that we will be sharing the collected data with other participants of the research. Will the social media channel itself cause a privacy violation? For example, a participant recruited via @Facebook may have consented to sharing their information contained on the specific page. However, the 'friends' of this person on @Facebook may well also have access to this data. In such cases, will @Facebook be able to protect the privacy of the participant? How anonymous is the person? The terms of service and end-user license agreements protect the researchers to some extent but do nothing to protect a participant under these circumstances. Traditional participant recruitment generally makes use of a 'push' strategy, where researchers know their potential targets. Even in the case of 'snowball sampling', participant groups tend to be controlled and recruiters have the ability to channel or restrict participant groups.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

By contrast, social media participant recruitment uses a 'pull' strategy. Potential participants are informed by means of an invitation and can talk to others, freely sharing information, without the researcher having any control or influence over this process. The responses to the participant invitation on social media are interactive rather than static, which may lead to unexpected outcomes. For example, if an invitation is sent out via @Facebook with participants posting responses and responding to posts made by others, subsequent posts may modify already-posted information. A tweet-based invitation may be re-tweeted with additional information added. Consequently, the information may become distorted, shared inaccurately and, as a result of interactivity, become transformed. And this inadvertent sharing may result in researchers losing control, with privacy and anonymity being compromised in the social media space.

The Notion of Privacy

One of the most relevant questions when recruiting via social media networks is the question of fairness to participants. In general, technology-mediated social media spaces are public and it is therefore reasonable to ask whether the space being used for research recruitment, participation and sharing information is perceived as private by its users. Specifically: are users of the space aware they are being observed and, if they are, would they still be happy to consent knowing their activities were being observed? In particular, are the inhabitants of the space expecting to be the subject of academic research? We argue that in many cases the participants are opting-in without much thought. If we take the example of a @Facebook-based invitation, the consequences of being observed do not occur to many participants.

In the case of an online research project, very few participants will truly be anonymous: IP addresses are usually traceable and online survey providers default to include them and 'tweets' may contain identifiers. Often, survey creators themselves are not familiar with the default settings on online survey tools such as *Qualtrics* or *SurveyMonkey*. In addition, the increasing inter-linking of social media and Big Data leads to ever-greater privacy concerns relating to research participants (Sula, 2016).

Informed consent is the fulcrum on which research balances, so that obtaining consent from adults online is crucial. In general, links from social networking sites to other, more reliable, sites could solve a number of privacy problems. For example: surveys using *Qualtrics* or *SurveyMonkey* can be accessed directly from advertisements in @Facebook, @Twitter or @LinkedIn; observable environments are easy to set up online; and in games environments, portals/tokens can be used. However, when the issue of informed consent relates to children or young people, there is no method of ensuring a participant's age or level of maturity. For example, are all @Facebook users really 15+ years old? It is often difficult to obtain parental consent and even harder to ensure participation of those who consented although it may be possible to access young people via their schools or their General Practitioner or some other relevant, formal relationship.

Security/Ownership of Data on Social Media

In Australia, Chapter 2 of the Australian Code (NHMRC, 2007) provides guidelines for ensuring security of data on social media sites. Relevant questions here include:

- How secure are your data in a social networking site? Can you access them as/when you need? Can anyone else access them at will?

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Can you depend on privacy / reliability claims from a social networking service, e.g. is your social network of choice always available?
- Can foreign governments access your data?

Social networking platforms are predominantly US-based, with real implications for research data (Hopewell, 2011). Any US-based data are subject to the Patriot Act, as well as to the less-well-known Foreign Intelligence and Surveillance Act. This means that data may be accessed by US Federal Law enforcement agencies, no matter who owns them. Australians storing data on US websites cannot claim protection under the Fourth Amendment of the US Constitution (which protects US citizens against unlawful search and seizure of property and information) because a 3rd-party provider stores their data. As well, inter-governmental treaties and agreements (e.g. the European Convention on Cybercrime) allow data stored in the US to be requisitioned by European law enforcement agencies without customer awareness, adding still further layers of government oversight.

The ownership of data on a social networking site is also crucial (Protalinksi, 2012). Who owns the data you create in a social networking site? Although it is not obvious to most users, @Facebook, in fact, claims the rights to any data collected from applications (including surveys) created. Jacquith (2009), explained this situation as follows. When anyone uses an application via @Facebook, it may ask for permission to access the @Facebook content and information, as well as content and information others have shared with you. While @Facebook claim to respect the privacy of the individual, the agreement with that application will control how it can use, store, transfer content and information. Therefore, when a @Facebook user publishes content or information using the public setting, that person is allowing everyone, including people who are not on @Facebook, to access and use the information or associate with the person.

This view that user data is the property of the social media site itself may well have been the root cause of recent scandals, such as that involving @Facebook's provision of user data to political campaigns – either directly for Barack Obama in 2008 and 2012, or via third-party Cambridge Analytica for Donald Trump in 2016 (see, for example, Pilkington & Michel, 2012; Urbain, 2018; Granville, 2018) – which have generated significant negative publicity for @Facebook during 2018. That political scandal has now widened into more general queries concerning @Facebook's access to and manipulation of other user data, such as health (Ostherr, 2018); and flow-on effects are starting to emerge for other major social media sites such as @Twitter and @Google's stable of products, as governments from Europe to the US to Canada and Australia question the data practices of the social media 'titans'(Browne, 2018; South China Morning Post, 2018). Under such conditions, one might well question whether Chapter 2 of the Australian Code can be actually followed.

Further complexities for Australian researchers exist in terms of the nation's strict privacy legislation and its implications for US-based social media sites. The 2014 changes to the federal *Privacy Act* mean that offshore data storage may place researchers in breach of certain Australian Privacy Principles (APPs 8 and/or 11.1 in this instance) of this Act, as those who send data overseas for storage are held responsible for any subsequent breaches or other dealings with those data, whether or not the researchers themselves have any influence over this process.

The research question was motivated and supported by the numerous questions raised in the previous sections: Is anonymity, privacy and security of participants recruited via social media (or linked methods) being protected in Australia? A study was undertaken in 2014 at University of Tasmania that

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

reviewed 18 months of ethics applications, at a time when social media use for recruitment, observation or research activity was still its infancy, with a view to understand researcher approaches and attitudes to this new environment. The study provided an early benchmark for ethical use of social media in research. The paper reports on the study findings and subsequently imbibing new knowledge from 2014 -18, adds an outlook.

RESEARCH DESIGN AND METHODOLOGY

For the initial phase of this project, we selected a set of ethics applications (n = 657) submitted over a period of 18 months to the University of Tasmania's Social Science Human Research Ethics Committee (SSHREC). Of these, 63 applications were found to have used social media as a tool to recruit applicants or, in a small number of cases, to gather data from participants once the project commenced.

Data Collection

The Project was formally reviewed / approved by the SSHREC's Deputy Chair to avoid a conflict of interest as one of the authors of this study was, at that time, the Chair of the SSHREC. Access to the server holding HRE data was formally approved in writing by the Research Operations Director and a research assistant (RA) with no other involvement in the study undertook the data retrieval and stripped out the fields relevant to the study, creating a fully-blinded dataset for the researchers. 657 applications submitted to the SSHREC between Jan 2013 – July 2014 were reviewed by the RA to identify those involving social media. Relevant, de-identified data from the 63 applications using social media to recruit participants / gather project data were then supplied by the RA to the researchers.

Data Coding

Data from the applications were coded into the following categories:

- Participant demographic 'class' (age, gender)
- Special selection issues (physical health, English ability)
- Privacy techniques (codes to replace login, secure storage, retention period)
- Recruitment approaches (invitation, snowball sampling, discussion boards)
- Social Media (@Facebook, @Instagram, @Twitter, @YouTube, @MySpace)
- Narratives for each element (if used).

Leximancer™: The Text Analytics Tool

Coded data were analysed using the *Leximancer*™ content analysis software to determine linkages and influences. This analytical software tool draws insights from qualitative data inputs by grouping related terms that are physically co-located in the text and are connected. The information is then displayed by means of a conceptual map which provides a 'bird's eye' view of the data and represents the main concepts and their relationships by forming thematic summaries (For example, see Figure 1).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Specifically, a concept is a group of related words (terms) that are linked together. The concepts begin with seed words (either automatically discovered by Leximancer™ or manually specified) for coding. When data are fed into Leximancer™ software, the analysis tool uses a process algorithm (known as knowledge discovery) and develops a list of concepts contained in the text and their relationships to one another. The outcome of this analysis process is in the form of a concept map. The (potentially iterative) learning process embedded in the tool evolves the concept seed definitions into a full thesaurus. Essentially, the tool automatically identifies concept seeds by looking for words which occur most frequently and groups them together. As the concept learning progresses, apart from adding highly relevant words to a concept, Leximancer™ may also add words that are negatively correlated with that concept (this rarely occurs). The aim of the concept learning process is to discover clusters of words which, taken together, maximise the relevance of other words.

A theme in this tool is a cluster of concepts bearing some common features or reflecting connectedness, as seen from their proximity on the visual concept map. The concept map is presented visually using coloured circles. However, the size of the circle (large/small) is irrelevant to the themes recognized. The circles are merely boundaries. Prevalence of a concept is determined by the number of concepts present in a theme, and this is given in the thematic report. The histogram bands produced for the thematic report are colour-coded (or heat mapped as hot or cold), to visually depict the prevalence of the theme. The colour is also carried through to the theme circle or boundary. The bar chart ranks the most appropriate themes relative to one another and, beneath that, the concepts which are visible within each theme are listed (for example, See Figure 2).

In *Leximancer*™, the term knowledge pathways refer to the strongest link between two concepts. As indicated in the visual analysis below, a black line on the concept cloud map indicates which other concepts might be bypassed to reach the concept themes. These pathways narrate stories that emerge from the textual analysis and focus on indirect connections between the concepts presented on the map. In the visual analysis, we present these pathways using the concept cloud view and knowledge pathways. The concept cloud view indicates the indirect connections that lead to the concept themes (or all the connections that are bypassed but are present to reach the destination theme).

The links between themes and the knowledge pathways generated from the concept maps are presented in the analytical discussion.

DATA VISUALISATION AND ANALYTICAL DISCUSSION

The researchers input the qualitative data obtained from the ethics applications into Leximancer™ to generate concept maps and thematic summaries (Figure 1).

The thematic summary shown in Figure 1 presents the main concepts and thematic summaries generated that are relevant and connected to the overall topic. Figure 2 represents the ranked concept list in order of relevance to themes. This figure shows three major concepts (data, participants and study) which were clearly inter-connected to all major themes.

Based on this preliminary analysis, we then classified the data into three main *concept themes* which capture the thematic summaries. These summaries reflect how the privacy of participants was protected in the research studies for which they were recruited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Concept maps and thematic summaries

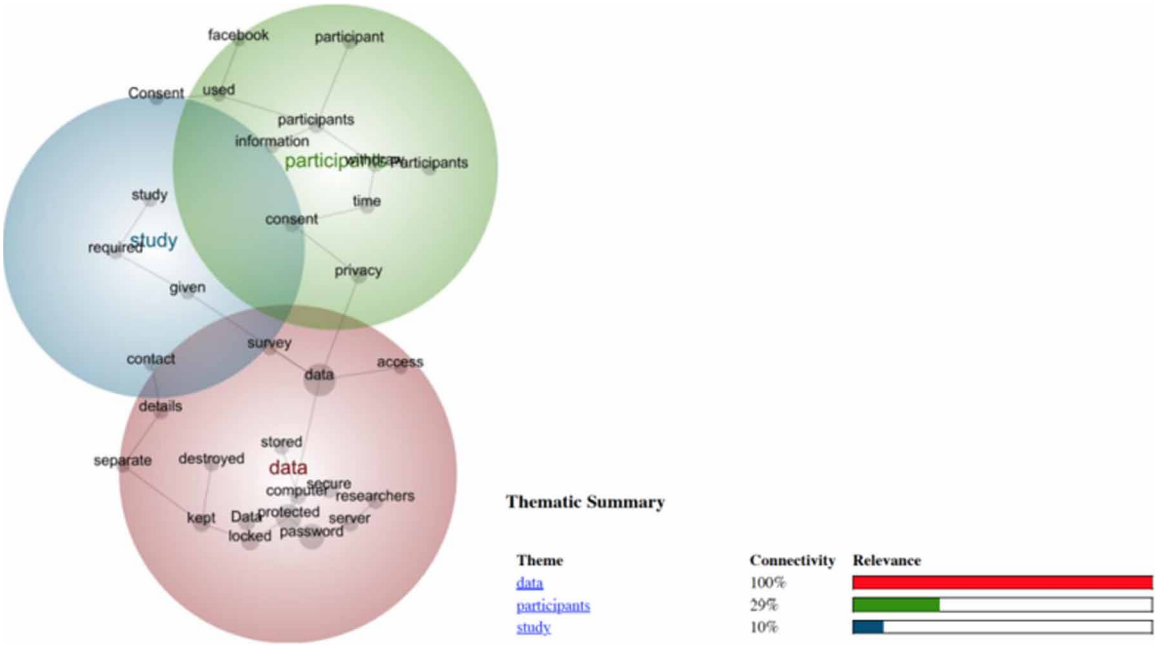
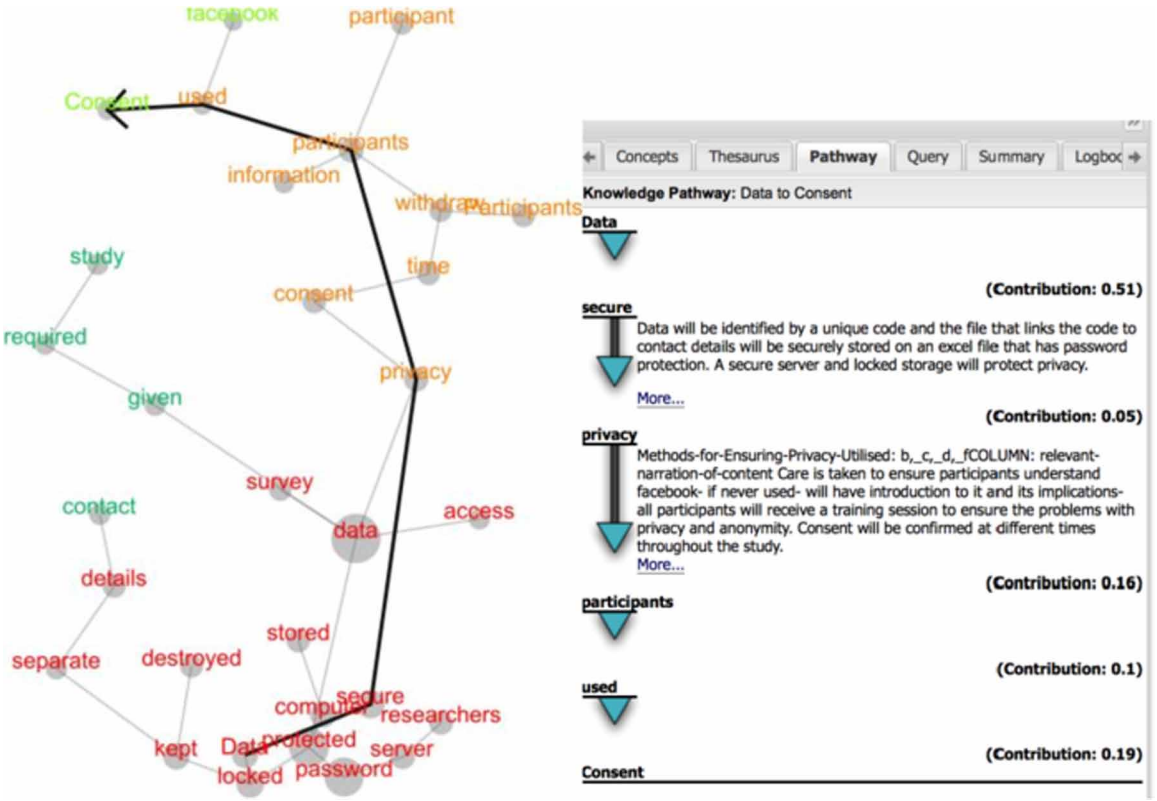


Figure 2. Ranked list of concepts and thematic summaries



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Concept Theme 1: Data

The theme named Data ranked the highest amongst the thematic summaries i.e. the prevalence of this theme and its linkage to other concepts emerged the strongest in the analysis (23% relevance). A key feature to note is that all research data is collected under conditions agreed to under HREC ethics approval. Data collection is based on informed consent during recruitment of research participants. This consent can be withdrawn (and therefore the participant's data can be removed) to protect privacy of participants until the publication of research involved. In practice, this meant that the collected data are extracted into word documents, anonymized, backed up into hard drives and locked in filing cabinets only accessible to investigators. Even data collected from blogs or posts (e.g. @Facebook) is protected, as the sharing of data ceases when a participant withdraws consent: even extracted/ anonymized data on word documents can be removed from the research data set. Paper copies of data are locked into a secure cabinet in the researchers' offices; electronic versions are kept on password protected computers exclusively. If any data are copied, there will be logs and records which can be deleted in case of withdrawal of consent.

Security of data is ensured by storage on secure servers within the research offices of the university, accessible only to research investigators listed on the ethics application. Data retained within the university are destroyed after 5 years (in the case of social research) and, in the case of projects where study outcomes had to be forwarded to partner organisations, only anonymized data were provided to government agencies (as per informed consent from the participants). In cases where a prize offer was provided to the participants for being part of the research the privacy and anonymity of data supplied were ensured by requiring participants wishing to be eligible for the prize to send a blank contact email to a designated 'prize' email address, thus separating the identity of the participant (their email) from the data provided for the study. In cases where questions in a survey could be regarded as intrusive, the survey design provided participants with the option not to answer. The data were, therefore, non-identifiable. Any access to the data is password-protected by University.

Clearly, this theme of data being anonymized and measured to protect the privacy of research participants emerged as a key theme in this research.

Concept Theme 2: Participants

While Data in particular protecting security and privacy is the key theme, the role and responsibility of participants is also highly relevant and prevalent (19%). Analysis of the Participant theme identified that the focus was on protecting privacy. The participants in the studies investigated were given the opportunity to protect their privacy through both direct and indirect methods. The participants themselves were encouraged to be involved in the process of protecting their own privacy. Prior to commencing data collection, research projects involving audio recordings, required participants consent to audio recording and, later, have the opportunity to review the transcript of their conversation. The notion of privacy was explicitly linked to eligibility criteria: if these criteria were not fulfilled, participation would not have been allowed. For example, the eligibility screening procedure warned participants that collection of data from those who, for example, disclosed illicit drug use (or any other activity which involved

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

researchers in a mandatory reporting process) would be placing themselves at risk (as required by the National Statement's Ch. 4.6).

Section 1.5 of the National Statement (NS) recommends that a summary of findings should always be provided to the participant group; and the results of the study showed that this was, indeed, occurring in all research project cases. Further, publications arising from project findings would not identify any individual participant: only de-identified and aggregated data were published.

For research projects in which @Facebook was used for recruitment, participants were provided information about @Facebook and its privacy implications; and a training session on protecting their privacy on social media (in this case, only @Facebook) was provided if they had not been regular users, or were new to @Facebook. Participants were informed that all data would be de-identified by the researchers. When @Facebook was used to invite participants, they were informed that participation was voluntary and that they could decline the invitation to take part in the project, without coercion, as required by section NS 2.2.

In addition to informed consent, participants were given the opportunity to use pseudonyms rather than their real names, to further increase their privacy. It was not necessary to obtain individual participant consent when accessing data relating to organisational social media: in this case, only a representative from the organisation was required to provide consent. However, even in this situation, privacy was protected as a software program anonymized social media data before the researchers accessed it.

Concept Theme 3: Consent

The *Consent* theme (11%) was the third most prevalent concept identified by the analysis. In particular, the notion of *informed consent* was highlighted in this theme. The research studies, analysed for this research, collected via electronic means a variety of information types, including: photos, recordings and written comments. Participants were therefore asked to provide consent in three different ways to reflect the nature of the data gathered. Consent forms for this data type included substantial information about privacy, rights and responsibilities. An information sheet was provided to all participants which detailed all ethical issues involved in sharing photos and other forms of information.

Further protection of participant privacy was provided by ensuring that 'screen names' (used in online forums) were only used when consent forms had been emailed to the researchers. If participants chose to withdraw from the study at a later stage, their responses would not be published, although their de-identified/anonymized responses would be used for analysis.

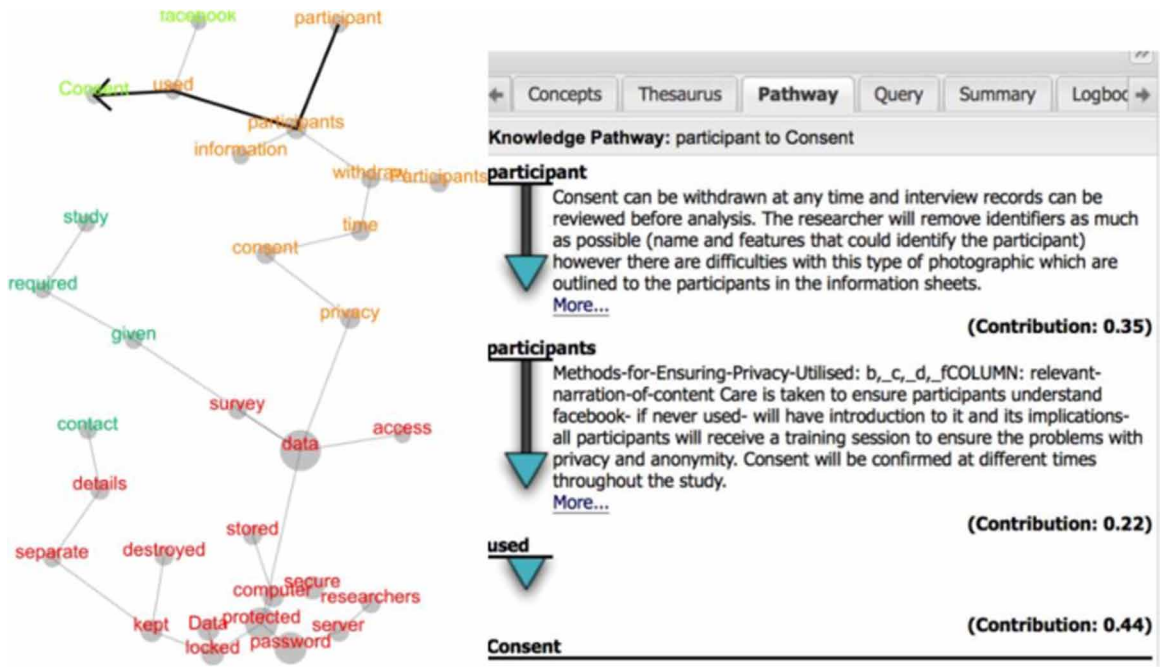
All participants had to sign consent forms (either electronically or physically) prior to participation and after being informed about privacy, rights and responsibilities. There was clear evidence in this emergent theme that informed consent was being obtained, ensuring privacy of participants.

Knowledge Pathways

In this section, we present the visualisation of Knowledge Pathways i.e. the strongest link between two concepts. The concept cloud view indicates the indirect connections that lead to the concept themes (or all the connections that are bypassed but are present to reach the destination theme).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. 'Data' to 'Participant' concept cloud and knowledge pathway



Pathway 1: Data to Participant

From the concept cloud shown in Figure 3, it can be seen that the black line connection between the strongest theme of 'Data' and that of 'Participants' is connected via secure and privacy concept pathways. As narrated in the knowledge pathway summary: the content can be scrolled past, privacy of participant's friends are preserved, cookies are deleted by using 'incognito' chrome session, logging out of profile in the presence of the participant. Data security is ensured by a unique code, together with the file that links the code to the contact details, which are securely stored in a password protected file.

The concepts of notions of privacy and security (both of data and participants) are apparent in this knowledge pathway connecting the two themes.

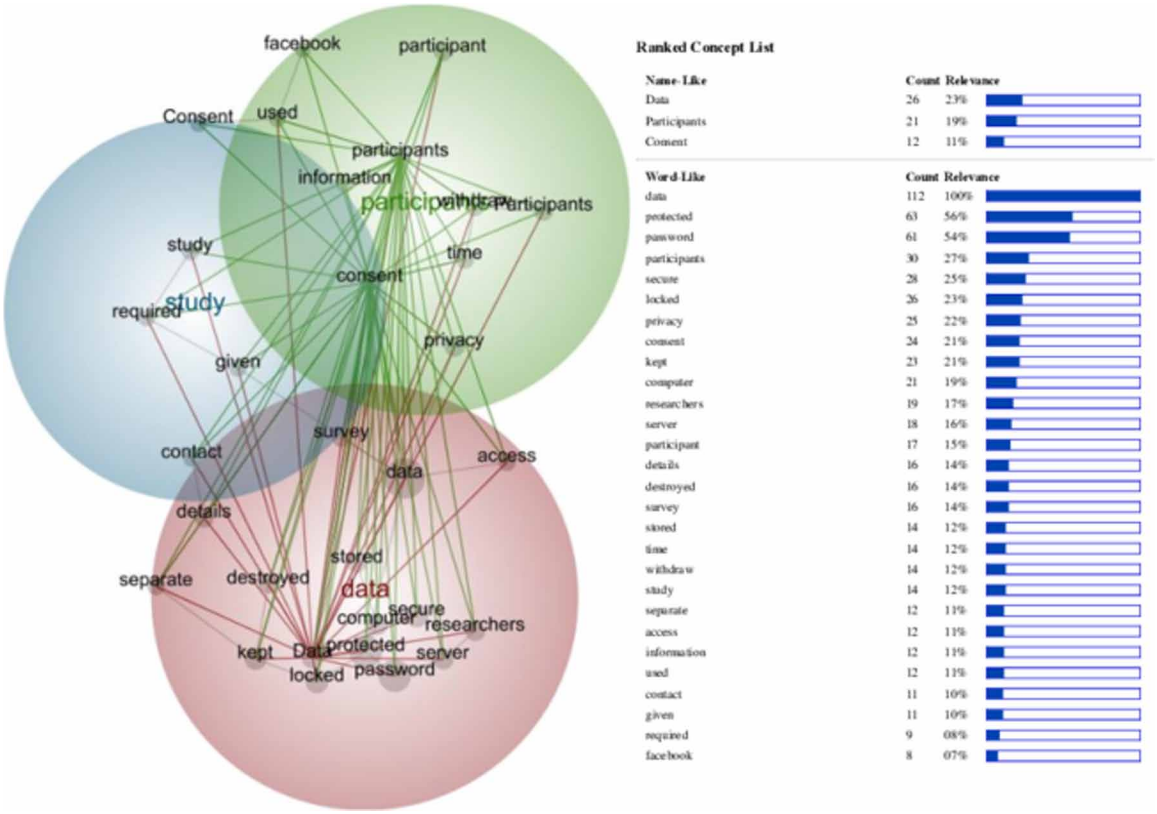
Pathway 2: Data to Consent

In this pathway, the notion of privacy between data and informed consent is examined. Security emerges as the strongest connection between the two themes. The notion of secure storage using passwords is a key element in facilitating informed consent. Specifically, participants are confident that the security of their data is ensured via a series of valid measures. At the same time, the researchers are ensuring privacy protection for participants by facilitating informed consent.

For example, see the commentary from *Knowledge Pathway: Privacy-Security*: '...relevant narration-of-content ... care is taken to ensure participants understand @Facebook – if ever used – will have an introduction to it and its implications – all participants will receive a training session to ensure the problems with privacy and anonymity...'

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. 'Data' to 'Consent' concept cloud and knowledge pathway



Therefore, there is a strong connection between enabling informed consent and protecting the data collected.

Pathway 3: Participant to Consent

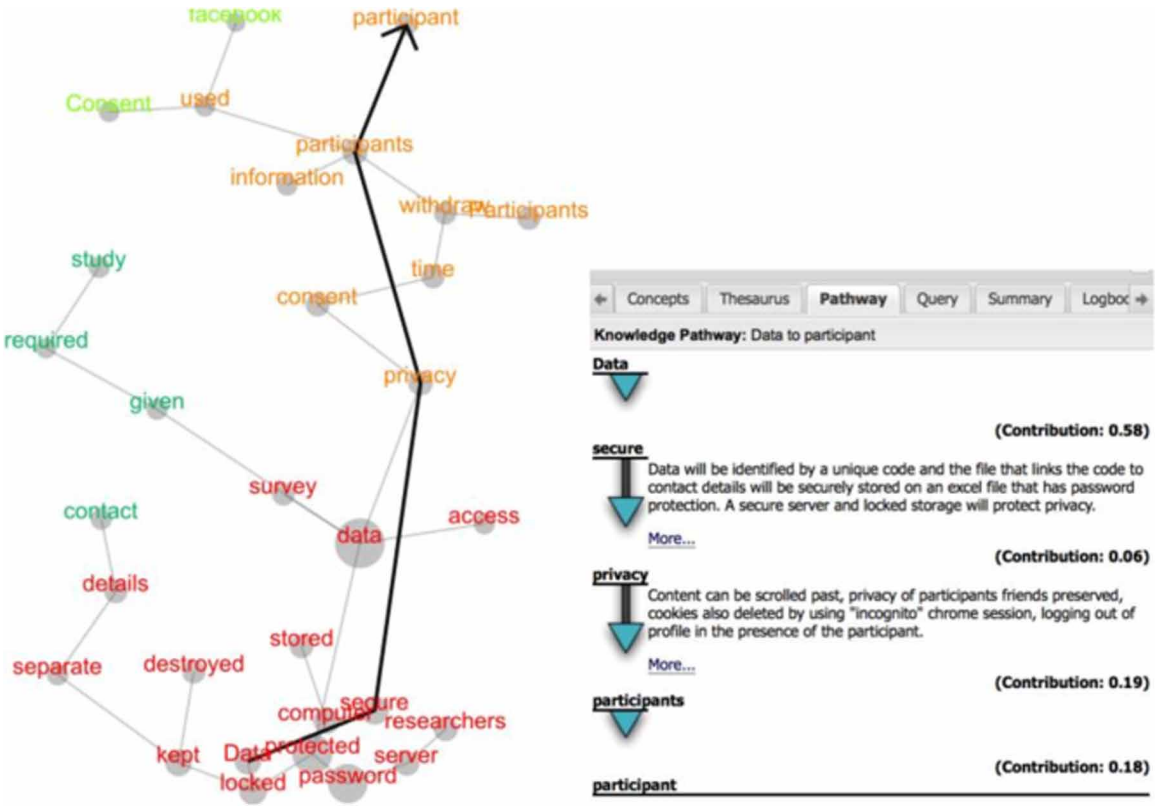
The shortest pathway seems to exist between the participant and the consent themes, predominantly because consent can be withdrawn at any time by participants. The researchers also ensure that identifiers are removed, to the greatest extent possible. Researchers also ensure participants are well informed, trained and made responsible for their own consent. The emerging story here is that participants are able to be confident that their privacy is protected, and that they can withdraw consent at any point in time.

In each of the visual analyses and summaries, evidence of sound and ethical practices using a variety of approaches for the protection of data anonymity, security, and privacy was reflected. The approaches included:

- Informed Consent – which can be withdrawn up to publishing date of data – used exclusively;
- Interest generation - @Facebook was used to generate interest in the study. Applicants were given a separate online link to the actual survey for identification and recruitment; and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. 'Participant' to 'Consent' concept and cloud knowledge pathway



- Online Surveys – using the @Facebook Messenger application and a specialised 'app' allowed tailored conversations between specific users. (No @Facebook-based surveys were included in the 63 applications reviewed)

As compared to a public post on @Facebook, a message and/or online survey via '@Facebook Messenger' can be considered personal for privacy purposes. In addition to the protection @Facebook Messenger offers, participants in this study were requested to remove the transcript of the @Facebook Messenger based online survey, after the responses were collected, to ensure that the privacy is fully protected.

All studies analyse the treated data (electronic and manual) in the same way. Data were transcribed, de-identified and retained securely for 5 years, after which they are destroyed. During this period, it was only possible for the researchers to re-identify data (for example in the case of a withdrawal of consent).

Addendum and Limitations

It should be noted that the ethics applications analysed as data for this research project were dated from early 2013 to mid-2014 (before and after the changes to the federal Privacy Law in Australia, which placed significant additional requirements on those storing data offshore). The number of ethics ap-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

plications involving social media recruitment approaches increased very rapidly (almost exponentially) over that period, supporting the general view of an upward trend in the use of social media for research recruitment and analysis.

At the time this research was conducted (in 2014), the system of ensuring anonymity, security and privacy of participants was apparently working for the Australian context. The privacy protection measures within the University combined with the federal Privacy Law were strong, and applicants were not relying on protection measures offered by @Facebook (the most popular social media site in Australia) or other social media providers. For example, arrangements were such that if a person is identified on @Facebook, they must re-identify themselves via an email link to be verified before recruitment can be completed (for this research study).

CONCLUSION AND OUTLOOK

This research project was conducted in 2014/15, at a time when privacy laws and statures were still evolving after the revisions to the federal Privacy Act. Since then, there has been a significantly increased uptake and saturation of smartphones in Australia. Most businesses have mobile applications and regularly undertake surveys, recruiting participants via social media. In this current scenario the question now arises for researchers (bound by the National Statement): would the same strategies apply for recruitment? No matter which approach is taken, researchers will need to ensure that data privacy and anonymity are being met in the ethical recruitment of participants: an issue which the @Facebook data privacy scandal makes immediate and urgent.

Contemporary researchers in Australia (such as Lunnay et al. (2015), conducted their research by recruiting participants through @Facebook. They cautioned qualitative researchers that ambiguity around the potential risks intrinsic to social media could hinder ethical conduct and discourage recruitment of participants through this medium. However, they also encouraged researchers that @Facebook recruitment is a practical research tool, adaptable for interacting with participants who already engage regularly with social media.

The North American research paper from Gelinas, Peirce et al. (2017) introduces a non-exceptionalism methodology for assessing social media recruitment, and these authors have identified 'respect for privacy' and 'investigator recruitment' as two key norms governing social media recruitment. According to this very recent, seminal work whether active or passive, social media recruitment has to be considered in the same way as traditional recruitment. The authors suggested that techniques should be normalised, while remaining sensitive to potentially novel aspects. They recommended mirroring physical-world methods explicitly, applying context-based ethical considerations and scrutiny; and then determining if social media recruitment differs from offline recruitment in ways that warrant a review. Their recommendations include two checklists to guide ethical recruitment of participants via social media. Allowing for some minor differences between the North American and Australian contexts, research using social media participation is today commonplace. Gelinas, Pierce et al. (2017) recommendations could be used as a guide where the recruitment of participants is achieved using social media.

We recommend that the recruitment of participants using social media extends both novel experience and convenience into the current Australian context still further, while the non-exceptionalist method – where traditional methods are mirrored when social media is used – could lead to a win-win situation. In

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Australia, the notion of privacy as guided by the National Statement is always translated into traditional recruitment approaches. Revisions to Chapter 3, published in 2018, includes further guidance on ethical online research approaches.

Our research findings made clear that the 'mirroring' recommended by Gelinas, Pierce et al. (2017) was already taking place, with participant privacy being well protected. Nonetheless, as social media-based recruitment becomes a norm for research, it is recommended that these checklists be used in addition to the National Health and Medical Research Council (NHMRC) of Australia guidelines, (both the National Statement and the Australian Code) contextualised to Australia in future. This method would ensure that the practical convenience of recruiting participants online is met with the notion of privacy still being protected – ensuring ethical research practices for the 21st century.

It would also be interesting to consider emerging research from other countries around the globe, and infer relevant insights into similar situation in Australia, and vice versa. We strongly recommend longitudinal and cross-country comparative studies in future to gain further insights.

ACKNOWLEDGMENT

The authors wish to thank the support of late Professor Ross Smith, for his feedback to enhance the quality of this paper. Dr Smith has rendered over 40 years to academia in Australia and this paper is his final contribution, before his death, in May 2019. The authors dedicate this paper to his eternal memory and invaluable contributions to academic research.

REFERENCES

AIATSIS. (2012). Guidelines for Ethical Research in Australian Indigenous Studies (2nd ed.). Retrieved from http://aiatsis.gov.au/research/ethical-research/guidelines-ethical-research-australian-indigenous-studies

Atkinson, J. (2013, March 31). Vintage Social Networking. Wrong Hands. Retrieved from http://wrong-hands1.wordpress.com/2013/03/31/vintage-social-networking/

ATNIX. (2013, January 21). ATNIX Australian Twitter News Index 2012 Roundup. QUT Centre for Creative Industry & Innovation. Retrieved from http://mappingonlinepublics.net/2013/01/21/atnix-australian-twitter-news-index-2012-round-up/

Blank, G. (2016). The digital divide among Twitter users and its implications for social research. *Social Science Computer Review*, *12*(October).

BMJ. (1996a). The Nuremberg Code 1947. British Medical Journal, 313(7070), 1148. Retrieved from http://www.cirp.org/library/ethics/nuremberg/

BMJ. (1996b). Declaration of Helsinki (1964). British Medical Journal, 313(7070). Retrieved from http://www.cirp.org/library/ethics/helsinki/

Browne. (2018). No doubt' other big tech firms could have similar data issues to @Facebook, top EU chief warns. CNBC. Retrieved from https://www.cnbc.com/2018/04/13/top-eu-chief-on-@Facebook-scandal-i-have-no-doubt-that-other-tech-firms-could-have-data-issues.html

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Buchanan, E. (2012). Social Media, Research, and Ethics: Challenges and Strategies. Rockefeller University Center for Clinical and Translational Science. Retrieved from http://www.uwstout.edu/faculty/buchanane/

Edison Research. (2012). The Social Habit. Retrieved from http://socialhabit.com/downloads/

Gelinas, L., Pierce, R., Winkler, S. I., Cohen, G., Fernandez Lynch, H., & Bierer, B. E. (2017). Using Social Media as a Research Recruitment Tool: Ethical issues and recommendations. *The American Journal of Bioethics*, *17*(3), 3–14. doi:10.1080/15265161.2016.1276644 PMID:28207365

Granville. (2018). @Facebook-Cambridge-Analytica. New York Times. Retrieved from https://www.nytimes.com/2018/03/19/technology/@Facebook-cambridge-analytica-explained.html

Henderson, M., Johnson, N. F., & Auld, G. (2013). Silences of ethical practice: Dilemmas for researchers using social media. *Educational Research and Evaluation*, *19*(6), 546–560. doi:10.1080/13803611.2013.805656

HHS.gov. (1979). The Belmont Report. US Dept. of Health & Human Services. Retrieved from https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

Hinchcliffe, D. (2011). Social business holds steady gap behind consumer social media. ZDNet. Retrieved from http://www.zdnet.com/blog/hinchcliffe/social-business-holds-steady-gap-behind-consumer-social-media/1695

Hopewell, L. (2011). The Pitfalls of Offshore Cloud. ZDNet.com.au. Retrieved from http://www.zdnet.com/the-pitfalls-of-offshore-cloud-1339308564/

Jacquith, M. (2009). @Facebook's bizarre definition of data ownership. Tempus Fugit blog. Retrieved from http://txfx.net/2009/02/26/@Facebooks-bizarre-definition-of-data-ownership/

Johnson, J. (2012). The Social Web Reveals Your Political Affiliation. The Inquisitr. Retrieved from http://www.inquisitr.com/273805/the-social-web-reveals-your-political-affiliation-infographic/#bF5m3jS8tepLhGh4.99

Lang, D. (2011). Where should you post your status? Breaking Copy blog. http://www.breakingcopy.com/social-media-flowchart-status

Lunnay, B., Borlagdan, J., McNaughton, D., & Ward, P. (2015). Ethical Use of Social Media to Facilitate Qualitative Research. *Qualitative Health Research*, *25*(1), 99–109. doi:10.1177/1049732314549031 PMID:25212856

MarketingProfs. (2012). Earned, Owned Media more trusted than Paid. Retrieved from http://www.marketingprofs.com/charts/2012/8943/earned-owned-media-more-trusted-than-paid

Meeker, M. (2017). Internet Trends 2017 – Code Conference. Kleiner Perkins Partners. Retrieved from https://www.scribd.com/document/349976485/mary-meeker-s-2017-internet-trends-report

Miller-Merrell. (2012). SEO basics for HR, recruiting and career pages. Blogging4Jobs. Retrieved from http://directemployers.org/2012/06/15/seo-bascis-for-hr-recruiting-and-career-pages/

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MIT. (2013). The Connected States of America, MIT Sensable City Lab & AT&T Research Labs. Retrieved from http://senseable.mit.edu/csa/

Mitchell, A., & Guskin, E. (2013). 'Twitter news consumers: young, mobile and educated. Pew Research Center Analysis. Retrieved from http://www.journalism.org/2013/11/04/twitter-news-consumers-young-mobile-and-educated/

NHMRC. (2003). Values and Ethics: Guidelines for Ethical Conduct in ATSI Health Research. Retrieved from https://www.nhmrc.gov.au/guidelines-publications/e52

NHMRC. (2007). Australian Code for Responsible Conduct in Research. Retrieved from https://www.nhmrc.gov.au/guidelines-publications/r39

NHMRC. (2016). Public consultation on Section 3. Retrieved from https://consultations.nhmrc.gov.au/public_consultations/ethical_conduct

Nielsen. (2012). Social Media Report 2012: Social Media comes of age. Retrieved from http://www.nielsen.com/us/en/newswire/2012/social-media-report-2012-social-media-comes-of-age.html

Ostherr. (2018). @Facebook knows a ton about your health — and plans to profit from it. Chicago Tribune. Retrieved from http://www.chicagotribune.com/news/opinion/commentary/ct-perspec-health-@Facebook-data-profit-0419-20180418-story,amp.html

MarkPack. (2015). Can you adjust data from Twitter to make it representative of the overall public's views? Retrieved from http://www.markpack.org.uk/134410/can-you-adjust-data-from-twitter-to-make-it-representative-of-the-overall-publics-views/

Pew Internet Research Group. (2012). Photos and Videos as social currency online. Retrieved from http://pewinternet.org/~/media//Files/Reports/2012/PIP_OnlineLifeinPictures.pdf

Pilkington & Michel. (2012). Obama digital data machine @Facebook election. The Guardian. Retrieved from https://www.theguardian.com/world/2012/feb/17/obama-digital-data-machine-@Facebook-election

Protalinksi, E. (2012). It's yours – until you upload it. In L. Dignan (Ed.), The Social Web: who owns your data? ZDNet Online Debate. Retrieved from http://www.zdnet.com/debate/the-social-web-who-owns-your-data/10087130/

Sensis. (2012). Yellow Social Media Report: what Australian businesses and people are doing with social media. Retrieved from http://about.sensis.com.au/IgnitionSuite/uploads/docs/FinalYellow_SocialMediaReport_digital_screen.pdf

Sensis. (2016). Sensis Social Media Report 2016: how Australian people and businesses are using social media. Retrieved from https://www.sensis.com.au/asset/PDFdirectory/Sensis_Social_Media_Report_2016.PDF

Sensis. (2017). Sensis Social Media report: Australians and Social Media. Retrieved from https://www.sensis.com.au/asset/PDFdirectory/Sensis-Social-Media-Report-2017.pdf

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

South China Morning Post. (2018). @Facebook data privacy scandal threatens fellow titans Google and Twitter as US Congress demands answers. Retrieved from http://www.scmp.com/news/world/united-states-canada/article/2139099/@Facebook-data-privacy-scandal-threatens-fellow

Sula, C. A. (2016). Research Ethics in an Age of Big Data. *Bulletin of the Assoc. for Science & Technology*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/bul2.2016.1720420207/full

University of Nevada Las Vegas (UNLV). (2016). History of Research Ethics. Retrieved from https://www.unlv.edu/research/ORI-HSR/history-ethics

Urbain. (2018, March 23). @Facebook as an election weapon, from Obama to Trump. Retrieved from https://phys.org/news/2018-03-@Facebook-election-weapon-obama-trump.html

This research was previously published in the International Journal of Virtual Communities and Social Networking (IJVCSN), 10(4); edited by Subhasish Dasgupta and Rohit Rampal; pages 16-32; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 78
Implications of Molecular Docking Assay for Bioremediation

Zarrin Basharat
Fatima Jinnah Women University, Pakistan

Monazza Bibi
Fatima Jinnah Women University, Pakistan

Azra Yasmin
Fatima Jinnah Women University, Pakistan

ABSTRACT

Bioremediation utilizes microbes to control environmental pollution, primarily through diverse enzymatic processes. With the incorporation of computation in biological experimentation, bioremediation has also been influenced by computational techniques. Molecular docking assay is one such pedestal of computational assisted bioremediation, which has been elaborated in this chapter. It helps in inferring whether the active site accommodate the pollutant molecules or not, depending on the stearic hindrance of the residues and nature of the active site pocket. The spotting of consequential active site residues and binding characteristics of compounds under study can conceivably be employed for site-directed mutagenic testing. From a vantage point, no one had expected such a remarkable usefulness of molecular docking assay for environmental research. Positive shades of low cost and efficiency, combined with eco-friendliness have made it a valuable method for analyzing biodegradative properties of enzymes responsible for pollution remediation.

DOI: 10.4018/978-1-7998-1204-3.ch078

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The world is facing an ever increasing menace of pollution due to industrialization and other anthropogenic activities. It is, therefore, at the foremost agenda of environmental scientists to find better, fast and eco-friendly solutions to mitigate pollution. Bioremediation is one such endeavor that exploits microbes to detoxify or degrade harmful contaminants in the environment. Bacterial enzymes are pollutant degradation workhorses that act to decontaminate environment. Molecular docking accelerates the procedure of choosing and improving enzymatic properties for a particular pollutant by quick computational scrutiny. The technique shows inordinate potential in its ability to predict enzyme adhesion and functionality in degrading or adsorbing the pollutant and presents a lucid picture of the relation and communication between pollutant and enzyme molecule. It helps in inferring whether the active site of certain enzymes accommodate the dye molecules or not, depending on the stearic hindrance of the residues and nature of the active site pocket. Protein-pollutant association is checked by assessing thousands of potential docking conformation poses. Picking poses that are reliable is an arduous task and one with lowest energy score is usually seen as the best match. Blind docking is attempted for proteins with unknown active site by entire protein surface scanning (Schnecke & Kuhn, 2000; Verdonk et al., 2005). Docking also provides information on mediator accommodation in the active site pocket to surmise about productive or non-favoured binding. The spotting of consequential active site residues and binding characteristics of compounds under study can conceivably be employed for site-directed mutagenic testing. These results then need to be tested in the lab experimentally and a critical evaluation of docking tools which gives better results might help define a set of rules and define a benchmark for such studies.

Molecular docking in remediation experienced a lag phase but now transitioned to alog phase in itsfrequency of application and attention of environmental scientists. It is now a noticeable area of research methodology assisting bioremediation studies. Bioinformatics boasts of numerous successful techniques developed for protein-ligand docking but shape complementarity and simulation approaches are most popular. In shape complementarity approach a matching procedure portrays the protein and the ligand as complementary surfaces (Meng, Shoichet, & Kuntz, 1992). In the simulative approach, interaction energies of ligand and protein are calculated (Feig, Onufriev, Lee, Im, Case, & Brooks, 2004). Knowledge derived information from docking assays for bioremediation is critical for a reliable evaluation of the state of the enzymatic micro-environment besides the transformations taking place inside enzyme, as well as for crafting strategies for sozotechnical measures aimed at environment clean up. This chapter discusses foundation of eco-analytical information for bioremediation, gained from molecular docking assay and provides information on some studies conducted until now. Finally the challenges and barriers to the molecular docking based bioremediation studies have also been analyzed and future perspective of such studies is also provided. The predictions are valuable for environmental chemists perusing metabolic intermediates, supervisory bodies/controllers endeavoring probable toxic products, microbiologists seeking to perceive microbial biodegradation, besides others having wide array of fascinations.

OVERVIEW

Molecular docking is a technique which calculates the preferred binding positioning of one molecule relative to the other, at a minimum energy conformation during complex formation (Lengauer & Rarey, 1996; Brooijmans & Kuntz, 2003). Docking can be carried out among protein and a small molecule,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

protein-protein, Protein-DNA, protein-RNA, protein-lipid, protein-sugar, polymer-ligand etc. where larger moiety is treated as a substrate and smaller one as the ligand. In case of protein-ligand interaction, which is the focus of this chapter, ligand generally fits inside cavity of the protein, which is usually predicted by a search algorithm (Westhead, Clark, & Murray, 1997; Gohlke, Hendlich, & Klebe, 2000; Brooijmans & Kuntz, 2003). This protein cavity turns active as a result of contact with substrates. It is formed of evolutionarily conserved residues and can thus, be traced on the basis of comparison with similar proteins (Lichtarge, Bourne, & Cohen, 1996). Many computational methods exist for identifying ligand-binding residues or active site prediction such as LIGSITE, SURFNET, CAST and PASS (Huang, 2009). After participation in the catalytic reaction, active site or catalytic site residue can be reused (Alberts et al., 2013). Some enzymes may have catalytic sites capable of binding effector molecule other than active site. These are called allosteric sites and occur mostly in proteins with more than one sub-unit and allow a range of interactions in addition to specific active site (Dedecker, 2000).

MECHANISM AND METHODS OF DOCKING

Protein-pollutant interaction study via docking is basically protein-ligand association modeling relative to bioremediation. Pollutants are treated flexible and allowed to adapt to a wide range of surface and pocket conformations. The complementary interaction region is also smaller and portrays reduced discrimination. The focus is therefore on the flexibility of the pollutant and modeling of weak associations accurately. 3D structure of the enzyme is first analyzed to identify active sites for fitting moiety of interest (Schechter, 2005). Experimentally obtained protein structures are the best source of information for protein-ligand docking studies. Availability of structural data information and steady increase in amount of publicly available structures with time, not only permits the investigation of novel protein-ligand docking parameters but also improve the understanding of already resolved as well as new targets. Structures relating a protein bound to diverse substrates or ligands could assist the knowledge discovery regarding ligand-induced conformation variations. New structures can be modeled using the already available structures (Basharat, 2015a). Standard computational methods are utilized forsolving protein structures based on the templates available in Protein Data Bank. More than 100,000 structures are currently present in the database and the number of sequences being deposited in NCBI alone is far greater than the number of these structures which means the number of protein sequences exceeds far greater than the number of structures and there is a huge gap, wealth of hidden information and demand for structure modeling. Steady progress has been made in this regard to model the protein structures based on available templates using the fold prediction and geometry calculation method called 'homology modeling'. *In silico* methods assist even the experimentally determined structures for geometry refining, estimation of binding energies, modeling the effect of mutations, predict the location of binding surfaces for new molecules.

Gap between sequence-to-structure ratio increases manifold in case of membrane proteins. This also signifies that the total numbers of residues are many times larger than the corresponding unique chains from the Protein Data Bank (Werner, Morris, Dastmalchi, & Church, 2012). *De novo* or *ab initio* modeling is employed in such a case for computational modeling of completely novel protein structures. These modeling approaches rely on physicochemical and thermodynamic properties of amino acid residues of the protein sequence and may not have direct linkage to the parameters of experimentally solved structures. Employing the scoring of energy function approach, it resorts to matching protein conformation at global minimum to least energy through an iterative procedure. *De novo* methods use knowledge based

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

scoring functions derived from experimental structures and their folds. Such methods are not template free *sensu stricto* and related to threading. *Ab initio* methods utilize force fields centered on first principles. This approach is template free and uses energy of the system for structure generation. Search space is large as all-atom model is used so this method works better in case of small sized proteins or peptides. The low-resolution structures obtained by computational approaches are further refined for docking studies through energy minimization and molecular dynamics (Werner et al., 2012). Protein structure assemblages are often implemented to obtain flexible models during protein-ligand docking. Structural conservation and catalytic site information is exploited for prediction of pollutant binding sites as catalytic residues and conformational changes in the active site region. Properly selected ensembles can improve docking performance considerably. Preprocessing of protein binding site ensembles may be carried out using dedicated webservers or databases like Astex Non-native Set (Verdonk, Mortenson, Hall, Hartshorn, & Murray, 2008), CoDNaS (Monzon, Juritz, Fornasari, & Parisi, 2013), Pocketome (An, Totrov, & Abagyan, 2005; Kufareva, Ilatovskiy, & Abagyan, 2012) and SIENA (Bietz & Rarey, 2016). The preparation of protein ensembles is quite intricate procedure due to unreliable annotation of many protein structures along with artifacts in structures. These impact the performance of ensemble employing applications. In case of missing or wrong information regarding binding site of a protein in the ensemble used for docking study, unintended bias may be introduced.

Ligand is usually obtained from a database source or drawn in a software or web interface (Exhaustive resource list available at http://www.click2drug.org/index.html) and saved. It is then processed and optimized before docking with the molecule. Missing residues and hydrogen atoms are added. Rotatable bonds are identified and partial atomic charges are calculated. As a ligand can move across six translational and rotational degrees of freedom, these movements are modeled. Initially both receptor and ligand molecules were treated as rigid bodies due to lack of computation power for solving the degrees of freedom. Maiden docking study of a protein and a small molecule (Kuntz, Blaney, Oatley, Langridge, & Ferrin, 1982) employed a negative receptor imageand six translational and rotational degrees of freedom. Spheres occupying the pockets and grooves on the receptor surface defined the prospective interaction sitesto clasp the ligand. Interactions were assessed by estimating the overlap as well as probablehydrogen bond formation between atoms in the protein-ligand complex. These types of rigid body calculations of protein-ligand studies have limitations as it does not account for induced fit ligand behavior (Brooijmans & Kuntz, 2013). With increment in computational power, algorithms and softwares were designed that could model interactions with one moiety as flexible and the other one as rigid (Leach, Shoichet, & Peishoff, 2006). Now, softwares exist for modeling the flexibility/degrees of freedom in both receptor molecule and ligand. The top results of active site accommodated ligands are judged on the basis of a statistical scoring function, which translates the unification energy into numerical value or the docking score (Leach et al., 2006). This is a critical step of docking, with correct prediction of the binding affinities/dissociation constant of a protein-ligand complex to separate true positives from false positives, a genuinely difficult task as its valuable outcome (Ambure & Roy, 2016; Yan & Wang, 2016). Complexes with least binding energies and maximum scores are then selected for further analysis. Calculation and scoring of interactions is a lot easier, less demanding and accurate in the gaseous phase as compared to that in the water phase due to small window of activity (Tirado-Rives & Jorgensen, 2006).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DOCKING ALGORITHMS AND SOFTWARES

For successful protein–ligand docking, speed and accuracy are the prime parameters. Development (Connolly, 1983) and usage of molecular surface calculations instead of explicit atom multi-configurations in docking have reduced the number of complexes and increased the speed of calculation (Jiang & Kim, 1991; Norel, Lin, Wolfson, & Nussinov, 1994). Ensuing this, grids to store physico-chemical properties of the receptor were integrated into docking procedures. Selected probe groups used were amino, methyl, and hydroxyl groups. The grids presented polar/nonpolar type of interactions, commonly preferred at each position in the binding site but presented difficulty in calculating energy. Later, tweaks in the algorithms led to hydrogen bonding, energy parameter and sophisticated scoring function calculation. This led to realistic scoring and the assessment of ligand flexibility in the active site during docking procedures. Ligand flexibility is usually calculated using systematic, stochastic, and deterministic search algorithms. Several algorithms use a sum of these approaches. Systematic search algorithms (Westhead et al., 1997) use grid values to calculate each degree of freedom with number of evaluations corresponding formal degrees of freedom. Cessation conditions are inserted to thwart the algorithm from sampling the space, identified to cause erroneous complex formation. Stochastic search algorithms check one degree of freedom at one time. Uncertainty of binding is rectified through independent runs. Examples include Monte Carlo methods and evolutionary algorithms. In deterministic searches, initial state leads to the next state, which is either equal or lower in energy to the previous state. These algorithms get trapped at local minima as they are unable to navigate across barriers. Different tactics are used to escalate their capacity to cross barriers or decrease the height of energy barriers. Examples include popular energy minimization and molecular dynamics simulation approaches. Other examples include Tabu search and Incremental construction etc. Some of these algorithms (Monte Carlo, Molecular dynamics and Genetic algorithms) are also applicable to receptor flexibility calculations (Brooijmans & Kuntz, 2013). For validation of the binding accuracy prediction of a protein with ligand, apart from the conventional root mean square deviation criteria, first principle, semi-empirical, empirical and knowledge-based methods exist and various softwares use these for calculating binding mode and affinities through estimation of parameters as van der Waals, electrostatic, hydrogen bond interactions etc (Kirchmair et al., 2012; Sliwoski, Kothiwale, Meiler, & Lowe, 2014). Classification and cataloguing of atom types, pairing this information with number of atom types and joint occurrence of specific class of atoms is essential part of knowledge based scoring functions. Free energy perturbation and thermodynamics incorporation are considered the high end procedures but computationally exhaustive. Molecular mechanics with Poisson Boltzmann surface area calculationapproach is also a recent popular method for free energy estimation of protein-ligand binding affinity (Weis, Katebzadeh, Söderhjelm, Nilsson, & Ryde, 2006; Liu, Liu, Zhu, Zhang, He, & Zhang, 2016). These approaches calculate the intermolecular, electrostatic, thermodynamic, stearic and solvent related forces or a combination of these forces. Algorithms/softwares for docking validation implement various amalgams of these forces for sieving out best results as neither of them is perfect in terms of precisionand broad applicability. Consensus scoring technique boosts the prospects of discovering accurately bound models by coalescing scores from several scoring functions (Huang, Grinter, & Zou, 2010).

More than fifty softwares are available for docking till to date, each varying in its approach or parameters but taking into account both these differing requirements. Speed of the fundamental steps has augmented due to availability of higher clock processing of modern computers along with the use of scoring grids. Latest developments in parallel computing domain have also enhanced the performance

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of docking algorithms and softwares considerably. Most popular of them are AutoDock (Huey, Morris, Olson, & Goodsell, 2007), Molegro Virtual Docker (Thomsen & Christensen, 2006), EADock (Grosdidier, Zoete, & Michielin, 2007), GOLD (Verdonk, Cole, Hartshorn, Murray, & Taylor, 2003; Joy, Nair, Hariharan, & Pillai, 2006), Flexx (Rarey, Kramer, Lengauer & Klebe, 1996; Kramer, Rarey, & Lengauer, 1999) etc. There are also dedicated web services available for docking like Molecular Docking Server (Bikadi & Hazai, 2009), SwissDock (Grosdidier, Zoete, & Michielin, 2011), PatchDock (Schneidman-Duhovny, Inbar, Nussinov, & Wolfson, 2005) etc for calculating geometry, site and energy of ligands interacting with proteins. The 3D pose of the bound pollutant-protein complex can be downloaded and visualized using different visualizing tools like PyMol, Rasmol etc which could help in inference of the best fit of ligand.

An increasing interest in human and animal enzyme interaction via molecular docking studies with pollutants is also of prime interest to environmental researchers working with occupational health hazard chemicals focused on bioaccumulation or toxicity property profiling of pollutants in human or animal bodies (Xie, Lü, & Chen, 2013; Zhou, Li, & Chen, 2013a; Liu, Chen, Jiang, & Song, 2014; Wang, Zhang, Cao, & Tang, 2014; Teng, Wang, Zou, Huang, & Du, 2015; Tripathi, Yasir, Gurjar, Bose, Dubey, & Shrivastava, 2015), Cytochrome P450 (Shimada et al., 2013; Zhou et al., 2013b; Shao, Chen, Li, & Zhao, 2014; Chushak, Chapleau, Frey, Mauzy, & Gearhart, 2015) and other xenobiotic degrading enzymes are of particular interest. Nanoparticle and polymer hazard has also been investigated through molecular docking assay for human and eco-toxicity bioindicator species (Lai et al, 2012; Hosseinzadeh, Moosavi-Movahedi, & Ghourchian, 2014; Liu et al., 2014; Yang, Hao, Li, Wei, Wang, & Liu, 2014; Ranjan, Dasgupta, Chinnappan, Ramalingam, & Kumar, 2015; Turabekova, Rasulev, Theodore, Jackman, Leszczynska, & Leszczynski, 2014). These are not being discussed in detail in this chapter as our focus is on bioremediation, and microbial enzyme docking investigation is best suited for such studies. Although, many of the enzymes in human and animal bodies also break down/detoxify the harmful pollutants but for environmental cleanup purpose, microbial enzymes are preferred owing to ease of availability and purification. Initiation and advancement of docking studies along the bioremediation line of research has been attempted by numerous researchers in the recent years (Table 1).

From the increasing number of studies including molecular docking for binding between pollutant and enzyme leading to degradation, it is evident that docking of protein-ligand is invaluable for studying the interactions of substrate and pollutants in bioremediation.

CASE STUDY: IBUPROFEN-COA LIGASE DOCKING

Sphyngomonas Ibu-2 has been reported to degrade ibuprofen [2-(4-Isobutylphenyl-Propionic acid)], an NSAID with an annual production of approximately 10 kilotonnes (Baumann & Baxendale, 2016), into catechols under aerobic conditions. CoA-ligase of *Sphyngomonas* sp. binds the said pharmaceutical and cleaves the acid side chain to yield isobutylcatechol metabolite (Murdoch & Hay, 2013). However, no report on the interaction of both these moieties exists. Here we describe a step by step method for protein-ligand study of this pollutant-enzyme interaction. The analysis shows that the CoA-ligase is capable of forming versatile conformations upon interaction with ibuprofen ligand. Modeling and abstraction played a noteworthy role in perceiving protein-ligand associations using computational *modus operandi*. Protein structural modeling and docking are helpful in inferring residues critical for interaction which might be mutated to obtain results of our interest. Detailed protocol and results are described as below:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Comprehensive list of microbe-pollutant interaction studies carried out using molecular docking assay

Organism	Enzyme	Pollutant	Docking software	Reference
Pseudomonas sp.	Naphthalene dioxygenase	Indole, Naphthalene, Biphenyl	AutoDock program	Carredano, Karlsson, Kauppi, Choudhury, Parales, Parales, Lee, Gibson, Eklund, & Ramaswamy, 2000
Burkholderia sp.	Biphenyl dioxygenase	2,3′-Chlorobiphenyl, 3,4′-Chlorobiphenyl	Sybyl	Zielinski, Kahl, Hecht, & Hofer, 2003
Sphingomonas sp.	Naphthalene dioxygenase	Polyaromatic hydrocarbons (indole, naphthalene, biphenyl)	GRAMM	Librando & Forte, 2005
Rhizobium sp.	Haloalkanoic dioxygenase	Chloropropionic acid	AutoDock	Hamid, Hamid, Haziyamin, Wahab, & Huyop, 2006
Sinorhizobium meliloti, Pseudomonas diminuta	Amidohydrolase superfamily	Hydantoin, Paraoxon	DOCK	Hermann, Ghanem, Li, Raushel, Irwin, & Shoichet, 2006
Trametes versicolor, Bacillus subtilis	Laccase	Anthracene, Phenanthrene, Thiodicarb, Malathion, Captan, Atrazine, Remazol Red B, Chlorophenol	GOLD	Suresh, Kumar, Kumar, & Singh, 2008
Caldariomyces fumago	Chloroperoxidase	Limonene	GOLD	Aguila, Vazquez-Duhalt, Tinoco, Rivera, Pecchi, & Alderete, 2008
Pseudomonas sp.	Monooxygenase	Aromatic hydrocarbon	PyMol	Notomista, Cafaro, Bozza, & Di Donato, 2008
Coriolus versicolor, Pleurotus ostreatus, Pseudomonas sp.	Laccase, manganese peroxidase, lignin peroxidase	Polyaromatic hydrocarbon	FRED	Arun, Raja, Arthi, Ananthi, Kumar, & Eyini, 2008
Comamonas testosteroni	Biphenyl dioxygenase	Polychlorinated biphenyls	AutoDock 4.0	Baig & Manickam, 2009
Bacillus subtilis	Azoreductase	Azo dyes	PatchDock	Ramanathan, Shanthi, & Sethumadhavan, 2009
Sphingomonas paucimobills	Haloalkane dehalogenase	1,2,3-trichloropropane	Affinity	Zhao, Zheng, Zhang, Chu, & Sun, 2009
Rigidoporus lignosus	Laccases	2,5-xylidine	AutoDock version 3.5	Cambria, Di Marino, Falconi, Garavaglia, & Cambria, 2010
Comamonas testosteroni	Biphenyl-2,3-dioxygenase	Polychlorinated biphenyls	AutoDock	Baig & Manickam, 2010
Sphingomonas sp. CHY1	Ring-hydroxylating dioxygenase	Polyaromatic hydrocarbons	AutodockVina	Librando & Pappalardo, 2011
Burkholderia xenovorans, Pandoraea pnomenusa	Biphenyl dioxygenase	DDT	AutoDock4.2	L'Abbée, Tu, Barriault, & Sylvestre, 2011
Enterococcus faecalis	Azoreductase	Methyl red	AutodockVina	Feng, Kweon, Xu, Cerniglia, & Chen, 2012
Aeromonas hydrophila	Laccase	Aromatic/phenolic compounds	AutoDock	Singh, Mani, & Chaudhary, 2012

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Continued

Organism	Enzyme	Pollutant	Docking software	Reference
Burkholderia cepacia.	Catechol 2, 3-dioxygenase	Petroleum hydrocarbons	Autodock 4.0	Ajao et al., 2012
Trametes versicolor	Laccase	Phenol	Molegro Virtual Docker	Zhang, Zeng, Zeng, Liu, Liu, Chen, Liu, Li, & Xie, 2012
Pycnoporus cinnabarinus	Laccase	Acid blue 74, Reactive black 5, Reactive blue 19	GOLD	Prasad et al., 2012
Brevundimonas diminuta	N-acyl-L-homoserine Lactone	Paraoxon	AutoDock	Zhan, Zhou, Guan, & Han, 2013
Rigidoporus lignosus	Laccase	2,5-Xylidine	AutoDock	Cambria et al., 2010
Geobacteraceae	Mn-Fe superoxide dismutase	Bisphenol-A, aminotriazole, ethylene-diurea	Molegro Virtual Docker	Emtiazi & Hedayatkha, 2014
Pandoraea pnomenusa	Biphenyl dehydrogenase	Polychlorinated biphenyls	HADDOCK	Piccoli, Musiani, & Giorgetti, 2014
Bacillus safensis	Laccase	Procion blue, Procion green, Procion red, Reactive yellow 86	GOLD	Singh, Sharma, Jacob, & Gakhar, 2014
E. coli	Azoreductase	Azo dyes	Autodockvina	Thakuria, Jungai, & Adhikari, 2015a
Rhodococcus sp. ustb-1	Naphthalene dioxygenase	Pyrene	Sybyl7.3	Jin, Yao, Zhang, Yu, Chen, Liu, Peng, & Choi, 2015
Pseudomonas putida	Phenol hydroxylase	Phenol, 2-chloro phenol	AutodockVina	Ray & Banarjee, 2015
Pseudomonas putida	Azoreductase	Azo dyes (Azobenzene,p-Aminoazobenzene, Amaranth dye, Methyl Orange, Sudan IV, Amidoschwarz, Congo Red, C.I. Food Yellow 3,)	Swiss Dock	Thakuria, Singha, Maisnam, & Adhikari, 2015b
Bacillus sp. B29	Azoreductase	azo dyes (orange I, orange II, orange G and acid red 88)	Molegro Virtual Docker 6.0	Dehghanian, Kay, & Kahrizi, 2015
Trametes versicolor	Laccase	Medicarpin	AutoDock 4.2	Martínez-Sotres, Rutiaga-Quiñones, Herrera-Bucio, Gallo, & López-Albarrán, 2015
Ceriporiopsis subvermispora	Laccase	Herbicide diuron	AutoDock4.2	Vieira, Marschalk, Biavatti, Lorscheider, Peralta, & Seixas, 2015
Bacillus sp.	Azoreductase	Methyl orange, Congo red	PatchDock	Basharat, 2015b
Bacillus sp.	Azoreductase	Azo dyes (orange G, acid red 88, Sudan I, orange I, and methyl red).	AutoDock	Haghshenas, Kay, Dehghanian, & Tavakol, 2015
Yersinia enterocolitica	Laccase	Aspirin, Diclofenac	GOLD	Singh, Rawat, Waseem, Gupta, Lynn, Nitin, Ramchiary, & Sharma, 2016

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Continued

Organism	Enzyme	Pollutant	Docking software	Reference
Aeromonas hydrophila	Laccase	Azo dyes(Joyfix Red, Remazol Red, Reactive Red, Reactive Yellow	Glide	Kumar, Shantkriti, Muruganandham, Murugesh, Rane, & Govindwar, 2016
Lysinibacillus sphaericus	Azoreductase	Azo dyes, JoyfixRed, Remazol Red, Reactive Red, Reactive Yellow	Glide	Kumar, Shantkriti,, Muruganandham, Murugesh, Rane, & Govindwar, 2016
Sulfobacillus acidophilus	Barbiturase	Atrazine, Simazine, Hexazinone	PatchDock	Basharat & Yasmin, 2016b

Step 1: Modeling, Validation, and Optimization of Enzyme Structure

The structure of *Sphingomonas* sp. CoA ligase (Accession no: ABK92273) was modeled using I-TASSER as previously described (Basharat & Yasmin, 2015a; Basharat & Yasmin, 2016a, Basharat, Messaoudi, Ruba, & Yasmin, 2016a). Ten templates were used for threading based model generation of the structure. Fatty acyl CoA synthetase of *Mycobacterium tuberculosis* (PDB ID: 3R44) was the top template with a structural similarity of more than 30%.Best generated model had aC-score of 0.44 (good quality model lies between -5 to 2 range). Estimated TM-score was 0.77±0.10 with a RMSD of 6.5±3.9Å. The structure was corrected for errors and missing residues. Fitting of intermolecular parameters to structural and thermodynamic data along with energy minimization was carried out until a gradient of 0.05 was achieved. The 3D protein structure (Figure 1) was then tested qualitatively using ERRAT (http://services.mbi.ucla.edu/ERRAT/) and Rampage (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php).

The structure had an overall quality factor of 78.19. Good resolution structures produce a value of over 95 and around 91 for 2.5-3 Å resolution structures. Since this is a homology model and equivalent to low resolution structure so the quality was considered fine. Structure passed the VERIFY3D (services.mbi.ucla.edu/Verify_3D) check as 90.51% of the residues had an average 3D-1D score of more than or equivalent to 0.2. According to the standard, more than 80% residues should have 3D-1D score more than or equivalent to 0.2. Ramachandran plot showed 79.2% of the residues were in favored region, 15.6% in allowed region and 5.1% in outliers.

Step 2: Ligand Procurement and Preparation

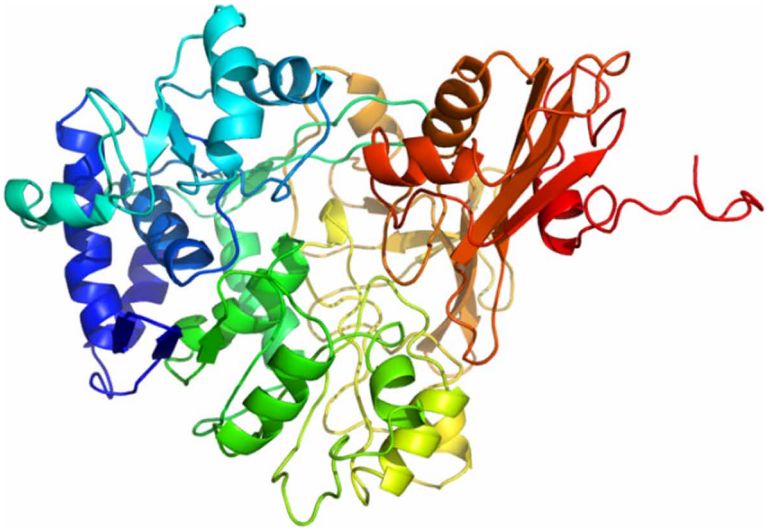
Ligand i.e. ibuprofen molecule was obtained from PubChem in 2D format and converted into 3D .pdb format using OpenBabel software. The ligand was optimized for pH 7, energy minimized using MMFF94 forcefield and saved for further analysis.

Step 3: Molecular Docking and Analysis

Docking was carried out using Molecular Operating Environmentsoftware.This software uses Monte Carlo simulated annealing process for docking, allowing conformational space sampling for ligand.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. 3D structure of the Sphyngomonas sp. ibuprofen-CoA ligase



Screening of everypossible binding site in the selected region of the target macromolecule was carried out. Docking interaction energy (U_{total}) of a given conformation is estimated from a set of energy grids centered in the macromolecule binding site using the formula parameters $U_{total} = U_{ele} + U_{vdw} + U_{ligand} + U_{solv}$ where U_{ele} is the electrostatic, U_{vdw} is the van der Waal interaction between protein-ligand complex. U_{ligand} is the conformation energy whereas U_{solv} represents the solvation energy. Among the generated poses, pose with lowest U_{total} is considered optimal. Ligase protein coordinates remained rigid during the docking process, while the flexible, mobile ligand relocated along the grid to trace most favorable binding orientation and conformation based on the interaction energy. The top ranking ten poses were retained in .mdb database format. These were then categorized according to increasing S-value i.e. final score indicating binding free energy. The complex with least S value was studied to evaluate the interactions of ligand with the active site residues of the ligase. Bonding interactions were evaluated via ligX option of Molecular Operating Environment. A value of -7.6335 was obtained for the optimal protein-ligand complex. 3D and 2D plot of the docked complex are shown in Figure 2 and Figure 3 respectively.

PERSPECTIVE

With increasing computation power, improved hardware and better simulations, we are achieving better insights and accurate prediction nearing natural chemical accuracy of interactions. This type of computational experimentation can help in unearthing the role of these residues and compounds in substrate specificity and activity, which in turn can be used for designing improved mutants and recombinants for degradation of noxious compounds (Prasad, Vindal, Narayana, Ramakrishna, Kunal, & Srinivas, 2012). Computational assessment of mutated protein residues in the structure is an inexpensive approach to assess alteration in functionality (Basharat & Yasmin, 2016c; Basharat, Zeb, & Yasmin, 2016b). This type of analysis attempted previously for laccase enzyme of *Bacillus licheniformis* using site-directed and random mutagenesis (Koschorreck, Schmid, & Urlacher, 2009) and can be extended for other enzymes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. 3D view of the protein-ligand complex. Docked ibuprofen is clearly visible in the pocket of the CoA-ligase of Sphyngomonas sp

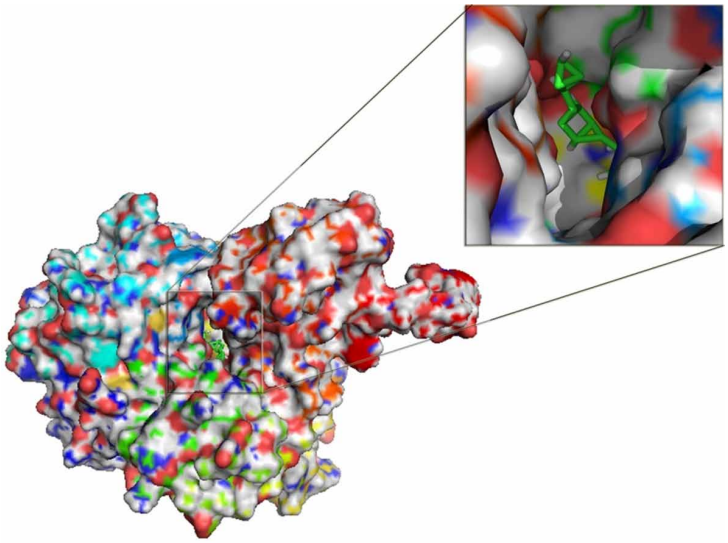
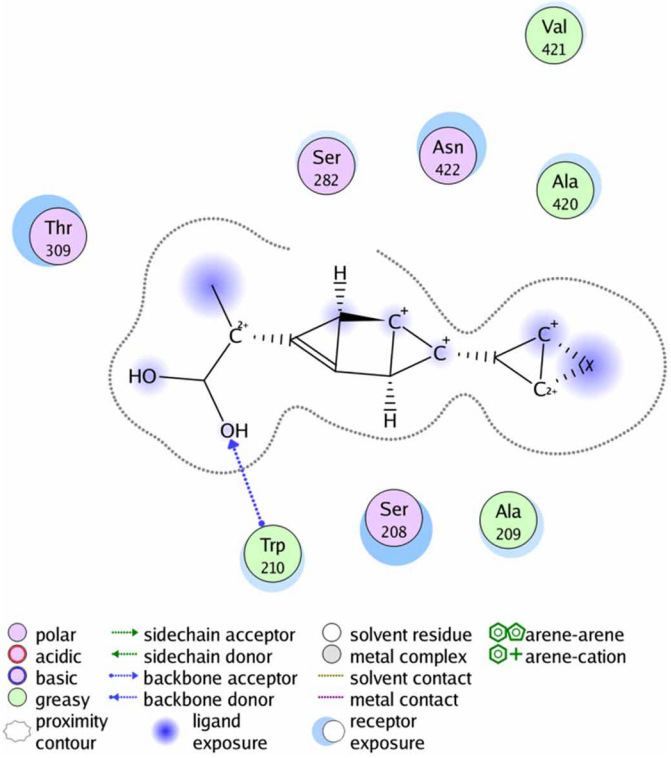


Figure 3. Molecular Operating Environment generated ligase CoA-ibuprofen interaction plot. It depicts 2D (flattened) spatial arrangement of ligand and CoA-ligase enzyme with respect to key interactions. The proximity contour shown by dashed lines. Trp210 donates electron to ibuprofen ligand OH group. The conformation of ligase-ibuprofen complex was energetically favored with the nucleophile and electrophile aptly positioned for catalysis



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

relevant to bioremediation. The existing algorithms and softwares are continually honed and perfected, as no one method functions fittingly for all protein systems.Paired occurrence of class of amino acids to more conjoined class of amino acids could help improve precision of scoring but is practically difficult to achieve due to scores of other undetermined parameters. Another challenge is the absence of a standard set of weighing coefficients for protein-ligand energy segment and forcefield scoring calculations. Presence of ions or co-factors in the pocket imposes parametrization difficulty in scoring as well. Evaluation of scoring functions is also a problem and each algorithm and software should be subject to CAPRI assessment for different datasets rather than simple root mean square deviation evaluation (for measuring quality of binding mode), before public release. Even though CAPRI (Janin, Henrick, Moult, Eyck, Sternberg, Vajda, Vakser, & Wodak, 2003; Méndez, Leplae, Lensink, & Wodak, 2005) exercises blind prediction feature that guarantees unbiased assessment but unrestricted usage of biological information like mutation dataand evolutionary analyses approach needs to be curtailed. Refined structural and binding datasets (including information of analogous pH and salt conditions) would be precious for parameter improvement. Systematic CASP-like ligand docking contests would do a good job of quality enhancement of algorithms.

Other challenges and barriers to the molecular docking include, but are not limited to issues like 'choosing the right approach', accuracy vs efficiency in assessing a particular algorithm, performance and prospective utility of a specific method over others in diverse situations etc. Throng of situations, factors, and data sets laid in front of the practitioner can confuse and the horde of pitfalls may render the docking approach less efficient. The shortcomings and technical traps of pollutant-enzyme docking methods could be avoided or reduced by knowledge based cautious approach which might make the practitioner aware and help make informed decisions which could limit probability of failure. Current scoring functions need to be trained rigorously on known and new datasets. In forcefield and knowledge based scoring methods, annexation of desolvation and entropic effects is indispensable for achieving precision.Scoring functions need to include fundamental compound measurements. Examples include transfer energies (for solvation and charges), dipoles/quadrupoles (for charge models), and pKa values for both small molecules and proteins (for solvation and binding energy). These can be calculated independently with softwares like DELPHI (Smith, Witham, Sarkar, Zhang, Li, Li, & Alexov, 2012; Sarkar, Witham, Zhang, Zhenirovskyy, Rocchia, & Alexov, 2013) but need to be integrated into the softwares and webservers.Such measurements are important as accurate scoring functions demands rigorous physical chemistry measurements for a wide variety of chemotypes.

Researchers have concentrated on pollutant binding to the protein, based upon structural knowledge of a known enzyme with degradation properties (Ajao, Kannan, Yakubu, Umoh. & Ameh, 2012; Basharat & Yasmin, 2016b). Conversely, designing a protein to bind a particular pollutant is also of interest to environmental scientists. It is intended for creating protein-based biosensors for pollutants. Computational design of pollutant-binding proteins is a perplexing inverse assessment of docking and scoring procedures, where experimental affinities are usually fed back into the algorithm development for refinement of complexes (Smith et al., 2016). Such protein design also entails a scoring function to rank and choose best conformations from the list of generated conformations. This would help in choosing relevant and active proteins to bioremediate/degrade pollutant.

Comparison between experimental and theoretical docking is not straightforward in biological entities but advent of combinatorial methods, growth in experimentally solved structures, theoretical advances, and improved computation power has definitely played a role in the success of docking studies and hopefully, will continue to do so in the near future. Computational efforts, particularly in combination with

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

relevant chemical strategies couldfurther boost efficiency and quality of docking. Molecular docking has been in use for more than three decades now and will continue to evolve and benefit the research community, including environmental scientists.

CONCLUSION

We need innovative real-time analytical methods capable of monitoring variations of priority pollutants in both time and location, integrated simulative models for perceiving and improving the breakdown mechanism, ultimately leading to better formulation of bioremediation strategies. To discover the substrate overtness of 3D available protein structure or a homology modeled enzyme and for gaining insight into the catalytic site, important residues and mechanism of the pollutant remediation-comprehensive information of enzyme structure complexed with substrate/pollutant is obligatory. For this purpose, computational molecular docking analysis is a boon and has been previously attempted for several noxious pollutants with enzymes. Although, several research groups have attempted to add to the scanty information on biodegradation analysis using docking assays but there is a need for more studies reporting and reproducing the conformations of protein-bound pollutants. The acquired information could aid in designing enzyme mutants for better accommodation of dye pollutants, especially for targeting at-once for a whole class of pollutants.

REFERENCES

Aguila, S., Vazquez-Duhalt, R., Tinoco, R., Rivera, M., Pecchi, G., & Alderete, J. B. (2008). Stereoselective oxidation of R-(+)-limonene by chloroperoxidase from *Caldariomyces fumago. Green Chemistry*, *10*(6), 647–653. doi:10.1039/b719992a

Ajao, A. T., Kannan, M., Yakubu, S. E., Umoh, V. J., & Ameh, J. B. (2012). Homology modeling, simulation and molecular docking studies of catechol-2, 3-Dioxygenase from *Burkholderia cepacia:* Involved in degradation of Petroleum hydrocarbons. *Bioinformation*, *8*(18), 848–854. doi:10.6026/97320630008848 PMID:23144539

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., ... Walter, P. (2013). *Essential Cell Biology*. Garland Science.

Ambure, P., & Roy, K. (2016). Scoring functions in docking experiments. In *Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery* (pp. 54–98). IGI Global. doi:10.4018/978-1-5225-0115-2.ch003

An, J., Totrov, M., & Abagyan, R. (2005). Pocketome via comprehensive identification and classification of ligand binding envelopes. *Molecular & Cellular Proteomics*, *4*(6), 752–761. doi:10.1074/mcp.M400159-MCP200 PMID:15757999

Arun, A., Raja, P. P., Arthi, R., Ananthi, M., Kumar, K. S., & Eyini, M. (2008). Polycyclic aromatic hydrocarbons (PAHs) biodegradation by basidiomycetes fungi, *Pseudomonas* isolate, and their cocultures: Comparative *in vivo* and *in silico* approach. *Applied Biochemistry and Biotechnology*, *151*(2-3), 132–142. doi:10.100712010-008-8160-0 PMID:18975143

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Baig, M. S., & Manickam, N. (2010). Homology modeling and docking studies of *Comamonas testosteroni* B-356 biphenyl-2, 3-dioxygenase involved in degradation of polychlorinated biphenyls. *International Journal of Biological Macromolecules*, *46*(1), 47–53. doi:10.1016/j.ijbiomac.2009.10.014 PMID:19879892

Basharat, Z. (2015a). Computational sequence analysis and *in silico* modeling of a stripe rust resistance protein encoded by wheat TaHSC70 gene. *Journal of Computational Biology and Bioinformatics Research*, *7*(1), 1–7. doi:10.5897/JCBBR2013.082

Basharat, Z. (2015b). *In vitro and in silico analysis of enzymes from indigenous bacteria for remediation of azo dyes* (Unpublished M.Phil. dissertation). Fatima Jinnah Women University, Rawalpindi, Pakistan.

Basharat, Z., Messaoudi, A., Ruba, S., & Yasmin, A. (2016a). NQO1 rs1800566 polymorph is more prone to NOx induced lung injury: Endorsing deleterious functionality through informatics approach. *Gene*, *591*(1), 14–20. doi:10.1016/j.gene.2016.06.048 PMID:27349566

Basharat, Z., & Yasmin, A. (2015a). *In silico* assessment of phosphorylation and O-β-GlcNAcylation sites in human NPC1 protein critical for Ebola virus entry. *Infection, Genetics and Evolution*, *34*, 326–338. doi:10.1016/j.meegid.2015.06.001 PMID:26048414

Basharat, Z., & Yasmin, A. (2016a). *Understanding properties of the master effector of phage shock operon in Mycobacterium tuberculosis via bioinformatics approach*. DOI:. doi:10.1101/050047

Basharat, Z., & Yasmin, A. (2016b). Molecular docking of *Sulfobacillus acidophilus* barbiturase with s-triazine compounds. *PeerJ Preprints, 4*, e2070v1. Doi:10.7287/peerj.preprints.2070v1

Basharat, Z., & Yasmin, A. (2016c). Bioinformatic analysis of human Gαq Q209 missense mutations associated with *uveal melanoma*. In *13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)* (pp. 106-111). IEEE. DOI:10.1109/IBCAST.2016.7429863

Basharat, Z., Zeb, S., & Yasmin, A. (2016b). *Computational study of some amoebicidal phytochemicals against heat shock protein of Naegleria fowleri*. Gene Reports; doi:10.1016/j.genrep.2016.09.003

Baumann, M., & Baxendale, I. R. (2016). Continuous photochemistry: The flow synthesis of ibuprofen via a photo-Favorskii rearrangement. *Reaction Chemistry & Engineering*, *1*(2), 147–150. doi:10.1039/C5RE00037H

Bietz, S., & Rarey, M. (2016). SIENA: Efficient compilation of selective protein binding site ensembles. *Journal of Chemical Information and Modeling*, *56*(1), 248–259. doi:10.1021/acs.jcim.5b00588 PMID:26759067

Bikadi, Z., & Hazai, E. (2009). Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. *Journal of Cheminformatics*, *1*(1), 1. doi:10.1186/1758-2946-1-15 PMID:20150996

Brooijmans, N., & Kuntz, I. D. (2003). Molecular recognition and docking algorithms. *Annual Review of Biophysics and Biomolecular Structure*, *32*(1), 335–373. doi:10.1146/annurev.biophys.32.110601.142532 PMID:12574069

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Cambria, M. T., Di Marino, D., Falconi, M., Garavaglia, S., & Cambria, A. (2010). Docking simulation and competitive experiments validate the interaction between the 2, 5-xylidine inhibitor and *Rigidoporus lignosus* laccase. *Journal of Biomolecular Structure & Dynamics*, *27*(4), 501–509. doi:10.1080/07391102.2010.10507334 PMID:19916571

Carredano, E., Karlsson, A., Kauppi, B., Choudhury, D., Parales, R. E., Parales, J. V., ... Ramaswamy, S. (2000). Substrate binding site of naphthalene 1, 2-dioxygenase: Functional implications of indole binding. *Journal of Molecular Biology*, *296*(2), 701–712. doi:10.1006/jmbi.1999.3462 PMID:10669618

Chushak, Y. G., Chapleau, R. R., Frey, J. S., Mauzy, C. A., & Gearhart, J. M. (2015). Identifying potential protein targets for toluene using a molecular similarity search, *in silico* docking and *in vitro* validation. *Toxicological Reviews*, *4*(2), 519–526.

Connolly, M. L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science*, *221*(4612), 709–713. doi:10.1126cience.6879170 PMID:6879170

DeDecker, B. S. (2000). Allosteric drugs: Thinking outside the active-site box. *Chemistry & Biology*, *7*(5), R103–R107. doi:10.1016/S1074-5521(00)00115-0 PMID:10801477

Dehghanian, F., Kay, M., & Kahrizi, D. (2015). A novel recombinant AzrC protein proposed by molecular docking and *in silico* analyses to improve azo dyes binding affinity. *Gene*, *569*(2), 233–238. doi:10.1016/j.gene.2015.05.063 PMID:26026905

Emtiazi, G., & Hedayatkhah, A. (2015). Comparative modelling of 3D-structure of *Geobacter* sp. M21 (a metal reducing bacteria) Mn-Fe superoxide dismutase and its binding properties with bisphenol-A, aminotriazole and ethylene-diurea. *Iranian Journal of Environmental Technology*, *1*(1), 19–29.

Feig, M., Onufriev, A., Lee, M. S., Im, W., Case, D. A., & Brooks, C. L. (2004). Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *Journal of Computational Chemistry*, *25*(2), 265–284. doi:10.1002/jcc.10378 PMID:14648625

Feng, J., Kweon, O., Xu, H., Cerniglia, C. E., & Chen, H. (2012). Probing the NADH-and Methyl Red-binding site of a FMN-dependent azoreductase (AzoA) from *Enterococcus faecalis. Archives of Biochemistry and Biophysics*, *520*(2), 99–107. doi:10.1016/j.abb.2012.02.010 PMID:22387379

Gohlke, H., Hendlich, M., & Klebe, G. (2000). Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, *295*(2), 337–356. doi:10.1006/jmbi.1999.3371 PMID:10623530

Grosdidier, A., Zoete, V., & Michielin, O. (2007). EADock: Docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins: Structure, Function, and Bioinformatics*, *67*(4), 1010–1025. doi:10.1002/prot.21367 PMID:17380512

Grosdidier, A., Zoete, V., & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Research*, *39*(suppl 2), W270–W277. doi:10.1093/nar/gkr366 PMID:21624888

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Haghshenas, H., Kay, M., Dehghanian, F., & Tavakol, H. (2016). Molecular dynamics study of biodegradation of azo dyes via their interactions with AzrC azoreductase. *Journal of Biomolecular Structure & Dynamics*, *34*(3), 453–462. doi:10.1080/07391102.2015.1039585 PMID:26325128

Hamid, A. A. A., Hamid, T. A., Haziyamin, T., Wahab, R. A., & Huyop, F. (2015). Identification of functional residues essential for dehalogenation by the non-stereospecific α-haloalkanoic acid dehalogenase from Rhizobium sp. RC1. *Journal of Basic Microbiology*, *55*(3), 324–330. doi:10.1002/jobm.201300526 PMID:25727054

Hermann, J. C., Ghanem, E., Li, Y., Raushel, F. M., Irwin, J. J., & Shoichet, B. K. (2006). Predicting substrates by docking high-energy intermediates to enzyme structures. *Journal of the American Chemical Society*, *128*(49), 15882–15891. doi:10.1021/ja065860f PMID:17147401

Hosseinzadeh, R., Moosavi-Movahedi, A. A., & Ghourchian, H. (2014). Electrochemistry and molecular modeling of the hemoglobin–benzene interaction with a nanocrystalline mixed metal oxide. *RSC Advances*, *4*(90), 49128–49136. doi:10.1039/C4RA08018A

Huang, B. (2009). MetaPocket: A meta approach to improve protein ligand binding site prediction. *OMICS: A Journal of Integrative Biology*, *13*(4), 325–330. doi:10.1089/omi.2009.0045 PMID:19645590

Huang, S. Y., Grinter, S. Z., & Zou, X. (2010). Scoring functions and their evaluation methods for protein–ligand docking: Recent advances and future directions. *Physical Chemistry Chemical Physics*, *12*(40), 12899–12908. doi:10.1039/c0cp00151a PMID:20730182

Huey, R., Morris, G. M., Olson, A. J., & Goodsell, D. S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*, *28*(6), 1145–1152. doi:10.1002/jcc.20634 PMID:17274016

Janin, J., Henrick, K., Moult, J., Eyck, L. T., Sternberg, M. J., Vajda, S., ... Wodak, S. J. (2003). CAPRI: A critical assessment of predicted interactions. *Proteins: Structure, Function, and Bioinformatics*, *52*(1), 2–9. doi:10.1002/prot.10381 PMID:12784359

Jiang, F., & Kim, S. H. (1991). Soft docking: Matching of molecular surface cubes. *Journal of Molecular Biology*, *219*(1), 79–102. doi:10.1016/0022-2836(91)90859-5 PMID:2023263

Jin, J. N., Yao, J., Zhang, Q. Y., Yu, C., Chen, P., Liu, W. J., ... Choi, M. M. (2015). An integrated approach of bioassay and molecular docking to study the dihydroxylation mechanism of pyrene by naphthalene dioxygenase in *Rhodococcus* sp. ustb-1. *Chemosphere*, *128*, 307–313. doi:10.1016/j.chemosphere.2015.02.012 PMID:25747183

Joy, S., Nair, P. S., Hariharan, R., & Pillai, M. R. (2006). Detailed comparison of the protein-ligand docking efficiencies of GOLD, a commercial package and ArgusLab, a licensable freeware. *In Silico Biology*, *6*(6), 601–605. PMID:17518767

Kirchmair, J., Williamson, M. J., Tyzack, J. D., Tan, L., Bond, P. J., Bender, A., & Glen, R. C. (2012). Computational prediction of metabolism: Sites, products, SAR, P450 enzyme dynamics, and mechanisms. *Journal of Chemical Information and Modeling*, *52*(3), 617–648. doi:10.1021/ci200542m PMID:22339582

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Koschorreck, K., Schmid, R. D., & Urlacher, V. B. (2009). Improving the functional expression of a *Bacillus licheniformis* laccase by random and site-directed mutagenesis. *BMC Biotechnology*, *9*(1), 1. doi:10.1186/1472-6750-9-12 PMID:19236694

Kramer, B., Rarey, M., & Lengauer, T. (1999). Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. *Proteins: Structure, Function, and Bioinformatics*, *37*(2), 228–241. doi:10.1002/(SICI)1097-0134(19991101)37:2<228::AID-PROT8>3.0.CO;2-8 PMID:10584068

Kufareva, I., Ilatovskiy, A. V., & Abagyan, R. (2012). Pocketome: An encyclopedia of small-molecule binding sites in 4D. *Nucleic Acids Research*, *40*(D1), D535–D540. doi:10.1093/nar/gkr825 PMID:22080553

Kumar, S. S., Shantkriti, S., Muruganandham, T., Murugesh, E., Rane, N., & Govindwar, S. P. (2016). Bioinformatics aided microbial approach for bioremediation of wastewater containing textile dyes. *Ecological Informatics*, *31*, 112–121. doi:10.1016/j.ecoinf.2015.12.001

Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., & Ferrin, T. E. (1982). A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, *161*(2), 269–288. doi:10.1016/0022-2836(82)90153-X PMID:7154081

LAbbée, J. B., Tu, Y., Barriault, D., & Sylvestre, M. (2011). Insight into the metabolism of 1, 1, 1-trichloro-2, 2-bis (4-chlorophenyl) ethane (DDT) by biphenyl dioxygenases. *Archives of Biochemistry and Biophysics*, *516*(1), 35–44. doi:10.1016/j.abb.2011.09.016 PMID:22001737

Lai, L., Lin, C., Xu, Z. Q., Han, X. L., Tian, F. F., Mei, P., ... Liu, Y. (2012). Spectroscopic studies on the interactions between CdTe quantum dots coated with different ligands and human serum albumin. *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*, *97*, 366–376. doi:10.1016/j.saa.2012.06.025 PMID:22797377

Leach, A. R., Shoichet, B. K., & Peishoff, C. E. (2006). Prediction of protein-ligand interactions. Docking and scoring: Successes and gaps. *Journal of Medicinal Chemistry*, *49*(20), 5851–5855. doi:10.1021/jm060999m PMID:17004700

Lengauer, T., & Rarey, M. (1996). Computational methods for biomolecular docking. *Current Opinion in Structural Biology*, *6*(3), 402–406. doi:10.1016/S0959-440X(96)80061-3 PMID:8804827

Librando, V., & Forte, S. (2005). Computer evaluation of protein segments removal effects from naphthalene 1, 2-dioxygenase enzyme on polycyclic aromatic hydrocarbons interaction. *Biochemical Engineering Journal*, *27*(2), 161–166. doi:10.1016/j.bej.2005.08.011

Librando, V., & Pappalardo, M. (2011). Computational study on the interaction of a ring-hydroxylating dioxygenase from *Sphingomonas* CHY-1 with PAHs. *Journal of Molecular Graphics & Modelling*, *29*(7), 915–919. doi:10.1016/j.jmgm.2011.03.001 PMID:21441049

Lichtarge, O., Bourne, H. R., & Cohen, F. E. (1996). An evolutionary trace method defines binding surfaces common to protein families. *Journal of Molecular Biology*, *257*(2), 342–358. doi:10.1006/jmbi.1996.0167 PMID:8609628

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, X., Liu, J., Zhu, T., Zhang, L., He, X., & Zhang, J. Z. (2016). PBSA_E: A PBSA based free energy estimator for protein-ligand binding affinity. *Journal of Chemical Information and Modeling*, *56*(5), 854–861. doi:10.1021/acs.jcim.6b00001 PMID:27088302

Liu, Y., Chen, M., Jiang, L., & Song, L. (2014). New insight into molecular interaction of heavy metal pollutant—cadmium (II) with human serum albumin. *Environmental Science and Pollution Research International*, *21*(11), 6994–7005. doi:10.100711356-014-2610-8 PMID:24522399

Martínez-Sotres, C., Rutiaga-Quiñones, J. G., Herrera-Bucio, R., Gallo, M., & López-Albarrán, P. (2015). Molecular docking insights into the inhibition of laccase activity by medicarpin. *Wood Science and Technology*, *49*(4), 857–868. doi:10.100700226-015-0734-8

Méndez, R., Leplae, R., Lensink, M. F., & Wodak, S. J. (2005). Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. *Proteins: Structure, Function, and Bioinformatics*, *60*(2), 150–169. doi:10.1002/prot.20551 PMID:15981261

Meng, E. C., Shoichet, B. K., & Kuntz, I. D. (1992). Automated docking with grid-based energy evaluation. *Journal of Computational Chemistry*, *13*(4), 505–524. doi:10.1002/jcc.540130412

Monzon, A. M., Juritz, E., Fornasari, M. S., & Parisi, G. (2013). CoDNaS: A database of conformational diversity in the native state of proteins. *Bioinformatics (Oxford, England)*, *29*(19), 2512–2514. doi:10.1093/bioinformatics/btt405 PMID:23846747

Murdoch, R. W., & Hay, A. G. (2013). Genetic and chemical characterization of ibuprofen degradation by *Sphingomonas* Ibu-2. *Microbiology*, *159*(3), 621–632. doi:10.1099/mic.0.062273-0 PMID:23329679

Norel, R., Lin, S. L., Wolfson, H. J., & Nussinov, R. (1994). Shape complementarity at protein–protein interfaces. *Biopolymers*, *34*(7), 933–940. doi:10.1002/bip.360340711 PMID:8054472

Notomista, E., Cafaro, V., Bozza, G., & Di Donato, A. (2009). Molecular determinants of the regioselectivity of toluene/o-xylene monooxygenase from *Pseudomonas* sp. strain OX1. *Applied and Environmental Microbiology*, *75*(3), 823–836. doi:10.1128/AEM.01951-08 PMID:19074607

Piccoli, S., Musiani, F., & Giorgetti, A. (2014). Dynamic characterization and substrate binding of cis-2, 3-dihydrobiphenyl-2, 3-diol dehydrogenase—an enzyme used in bioremediation. *Journal of Molecular Modeling*, *20*(12), 1–6. doi:10.100700894-014-2531-y PMID:25433599

Prasad, N. K., Vindal, V., Narayana, S. L., Ramakrishna, V., Kunal, S. P., & Srinivas, M. (2012). *In silico* analysis of *Pycnoporus cinnabarinus* laccase active site with toxic industrial dyes. *Journal of Molecular Modeling*, *18*(5), 2013–2019. doi:10.100700894-011-1215-0 PMID:21877154

Ramanathan, K., Shanthi, V., & Sethumadhavan, R. (2009). *In silico* identification of catalytic residues in azobenzene reductase from *Bacillus subtilis* and its docking studies with azo dyes. *Interdisciplinary Sciences: Computational Life Sciences*, *1*(4), 290–297. PMID:20640807

Ranjan, S., Dasgupta, N., Chinnappan, S., Ramalingam, C., & Kumar, A. (2015). A novel approach to evaluate titanium dioxide nanoparticle–protein interaction through docking: An insight into mechanism of action. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, 1-7.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rarey, M., Kramer, B., Lengauer, T., & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*, *261*(3), 470–489. doi:10.1006/jmbi.1996.0477 PMID:8780787

Ray, S., & Banerjee, A. (2015). Molecular level biodegradation of phenol and its derivatives through dmp operon of *Pseudomonas putida:* A bio-molecular modeling and docking analysis. *Journal of Environmental Sciences (China)*, *36*, 144–151. doi:10.1016/j.jes.2015.03.035 PMID:26456616

Rivera-Hoyos, C. M., Morales-Álvarez, E. D., Poveda-Cuevas, S. A., Reyes-Guzmán, E. A., Poutou-Piñales, R. A., Reyes-Montaño, E. A., ... Cardozo-Bernal, Á. M. (2015). Computational analysis and low-scale constitutive expression of laccases synthetic genes GlLCC1 from *Ganoderma lucidum* and POXA 1B from *Pleurotus ostreatus* in *Pichia pastoris. PLoS ONE*, *10*(1), e0116524. doi:10.1371/journal.pone.0116524 PMID:25611746

Sarkar, S., Witham, S., Zhang, J., Zhenirovskyy, M., Rocchia, W., & Alexov, E. (2013). DelPhi Web Server: A comprehensive online suite for electrostatic calculations of biological macromolecules and their complexes. *Communications in Computational Physics*, *13*(01), 269–284. doi:10.4208/cicp.300611.201011s PMID:24683424

Schechter, I. (2005). Mapping of the active site of proteases in the 1960s and rational design of inhibitors/drugs in the 1990s. *Current Protein & Peptide Science*, *6*(6), 501–512. doi:10.2174/138920305774933286 PMID:16381600

Schnecke, V., & Kuhn, L. A. (2000). Virtual screening with solvation and ligand-induced complementarity. *Perspectives in Drug Discovery and Design*, *20*(1), 171–190. doi:10.1023/A:1008737207775

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*, *33*(Web Serversuppl 2), W363–W367. doi:10.1093/nar/gki481 PMID:15980490

Shao, J., Chen, J., Li, T., & Zhao, X. (2014). Spectroscopic and molecular docking studies of the *in vitro* interaction between puerarin and cytochrome P450. *Molecules (Basel, Switzerland)*, *19*(4), 4760–4769. doi:10.3390/molecules19044760 PMID:24743933

Shimada, T., Kim, D., Murayama, N., Tanaka, K., Takenaka, S., Nagy, L. D., ... Guengerich, F. P. (2013). Binding of diverse environmental chemicals with human cytochromes P450 2A13, 2A6, and 1B1 and enzyme inhibition. *Chemical Research in Toxicology*, *26*(4), 517–528. doi:10.1021/tx300492j PMID:23432429

Singh, D., Rawat, S., Waseem, M., Gupta, S., Lynn, A., Nitin, M., ... Sharma, K. K. (2016). Molecular modeling and simulation studies of recombinant laccase from *Yersinia enterocolitica* suggests significant role in the biotransformation of non-steroidal anti-inflammatory drugs. *Biochemical and Biophysical Research Communications*, *469*(2), 306–312. doi:10.1016/j.bbrc.2015.11.096 PMID:26631965

Singh, D., Sharma, K. K., Jacob, S., & Gakhar, S. K. (2014). Molecular docking of laccase protein from *Bacillus Safensis* DSKK5 Isolated from earthworm gut: A novel method to study dye decolorization potential. *Water, Air, and Soil Pollution*, *225*(11), 1–12. doi:10.100711270-014-2175-7

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Singh, V., Mani, I., & Chaudhary, D. K. (2012). Analysis of the multicopper oxidase gene regulatory network of *Aeromonas hydrophila. Systems and Synthetic Biology*, *6*(3-4), 51–59. doi:10.100711693-012-9096-9 PMID:24294339

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). Computational methods in drug discovery. *Pharmacological Reviews*, *66*(1), 334–395. doi:10.1124/pr.112.007336 PMID:24381236

Smith, N., Witham, S., Sarkar, S., Zhang, J., Li, L., Li, C., & Alexov, E. (2012). DelPhi web server v2: Incorporating atomic-style geometrical figures into the computational protocol. *Bioinformatics (Oxford, England)*, *28*(12), 1655–1657. doi:10.1093/bioinformatics/bts200 PMID:22531215

Smith, R. D., Damm-Ganamet, K. L., Dunbar, J. B. Jr, Ahmed, A., Chinnaswamy, K., Delproposto, J. E., ... Doyle, L. (2016). CSAR benchmark exercise 2013: Evaluation of results from a combined computational protein design, docking, and scoring/ranking challenge. *Journal of Chemical Information and Modeling*, *56*(6), 1022–1031. doi:10.1021/acs.jcim.5b00387 PMID:26419257

Suresh, P. S., Kumar, A., Kumar, R., & Singh, V. P. (2008). An *Insilco* approach to bioremediation: Laccase as a case study. *Journal of Molecular Graphics & Modelling*, *26*(5), 845–849. doi:10.1016/j.jmgm.2007.05.005 PMID:17606396

Teng, Y., Wang, X., Zou, L., Huang, M., & Du, X. (2015). Experimental and theoretical study on the binding of 2-mercaptothiazoline to bovine serum albumin. *Journal of Luminescence*, *161*, 14–19. doi:10.1016/j.jlumin.2014.12.046

Thakuria, B., Jungai, N., & Adhikari, S. (2015). Catalytic Site Prediction of azoreductase enzyme of *E. coli* with potentially important industrial dyes using molecular docking tools. *International Journal of Bioscience. Biochemistry and Bioinformatics*, *5*(2), 91.

Thakuria, B., Singha, C. J., Maisnam, P., & Adhikari, S. (2015). Functional and catalytic active sites prediction and docking analysis of azoreductase enzyme in *Pseudomonas putida* with a variety of commercially available azodyes. *African Journal of Biotechnology*, *14*(26), 2162–2169. doi:10.5897/AJB2015.14699

Thomsen, R., & Christensen, M. H. (2006). MolDock: A new technique for high-accuracy molecular docking. *Journal of Medicinal Chemistry*, *49*(11), 3315–3321. doi:10.1021/jm051197e PMID:16722650

Tirado-Rives, J., & Jorgensen, W. L. (2006). Contribution of conformer focusing to the uncertainty in predicting free energies for protein-ligand binding. *Journal of Medicinal Chemistry*, *49*(20), 5880–5884. doi:10.1021/jm060763i PMID:17004703

Tripathi, M. K., Yasir, M., Gurjar, V. S., Bose, P., Dubey, A., & Shrivastava, R. (2015). Insights from the molecular docking of hydrolytic products of methyl isocyanate (MIC) to inhibition of human immune proteins. *Interdisciplinary Sciences: Computational Life Sciences*, *7*(3), 287–294. PMID:26297312

Turabekova, M., Rasulev, B., Theodore, M., Jackman, J., Leszczynska, D., & Leszczynski, J. (2014). Immunotoxicity of nanoparticles: A computational study suggests that CNTs and C 60 fullerenes might be recognized as pathogens by Toll-like receptors. *Nanoscale*, *6*(7), 3488–3495. doi:10.1039/c3nr05772k PMID:24548972

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Verdonk, M. L., Chessari, G., Cole, J. C., Hartshorn, M. J., Murray, C. W., Nissink, J. W. M., ... Taylor, R. (2005). Modeling water molecules in protein-ligand docking using GOLD. *Journal of Medicinal Chemistry*, *48*(20), 6504–6515. doi:10.1021/jm050543p PMID:16190776

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., & Taylor, R. D. (2003). Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*, *52*(4), 609–623. doi:10.1002/prot.10465 PMID:12910460

Verdonk, M. L., Mortenson, P. N., Hall, R. J., Hartshorn, M. J., & Murray, C. W. (2008). Protein− ligand docking against non-native protein conformers. *Journal of Chemical Information and Modeling*, *48*(11), 2214–2225. doi:10.1021/ci8002254 PMID:18954138

Vieira, A. C., Marschalk, C., Biavatti, D. C., Lorscheider, C. A., Peralta, R. M., & Seixas, F. A. V. (2015). Modeling based structural insights into biodegradation of the herbicide diuron by laccase-1 from *Ceriporiopsis subvermispora*. *Bioinformation*, *11*(5), 224–228. doi:10.6026/97320630011224 PMID:26124565

Wang, Y. Q., Zhang, H. M., Cao, J., & Tang, B. P. (2014). Binding of a new bisphenol analogue, bisphenol S to bovine serum albumin and calf thymus DNA. *Journal of Photochemistry and Photobiology. B, Biology*, *138*, 182–190. doi:10.1016/j.jphotobiol.2014.06.002 PMID:24972352

Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I., & Ryde, U. (2006). Ligand affinities predicted with the MM/PBSA method: Dependence on the simulation method and the force field. *Journal of Medicinal Chemistry*, *49*(22), 6596–6606. doi:10.1021/jm0608210 PMID:17064078

Werner, T., Morris, M. B., Dastmalchi, S., & Church, W. B. (2012). Structural modelling and dynamics of proteins for insights into drug interactions. *Advanced Drug Delivery Reviews*, *64*(4), 323–343. doi:10.1016/j.addr.2011.11.011 PMID:22155026

Westhead, D. R., Clark, D. E., & Murray, C. W. (1997). A comparison of heuristic search algorithms for molecular docking. *Journal of Computer-Aided Molecular Design*, *11*(3), 209–228. doi:10.1023/A:1007934310264 PMID:9263849

Xie, X., Lü, W., & Chen, X. (2013). Binding of the endocrine disruptors 4-tert-octylphenol and 4-nonylphenol to human serum albumin. *Journal of Hazardous Materials*, *248*, 347–354. doi:10.1016/j.jhazmat.2013.01.036 PMID:23416871

Yan, Z., & Wang, J. (2016). Scoring Functions of Protein-Ligand Interactions. In *Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery* (pp. 220–245). IGI Global. doi:10.4018/978-1-5225-0115-2.ch009

Yang, B., Hao, F., Li, J., Wei, K., Wang, W., & Liu, R. (2014). Characterization of the binding of chrysoidine, an illegal food additive to bovine serum albumin. *Food and Chemical Toxicology*, *65*, 227–232. doi:10.1016/j.fct.2013.12.047 PMID:24394486

Zhan, D., Zhou, Z., Guan, S., & Han, W. (2013). The Effect of conformational variability of phosphotriesterase upon N-acyl-L-homoserine lactone and paraoxon binding: Insights from molecular dynamics studies. *Molecules (Basel, Switzerland)*, *18*(12), 15501–15518. doi:10.3390/molecules181215501 PMID:24352010

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zhang, Y., Zeng, Z., Zeng, G., Liu, X., Liu, Z., Chen, M., ... Xie, G. (2012). Effect of Triton X-100 on the removal of aqueous phenol by laccase analyzed with a combined approach of experiments and molecular docking. *Colloids and Surfaces. B, Biointerfaces*, *97*, 7–12. doi:10.1016/j.colsurfb.2012.04.001 PMID:22580478

Zhao, Y. S., Zheng, Q. C., Zhang, H. X., Chu, H. Y., & Sun, C. C. (2009). Interaction between 1, 2, 3-trichloropropane and haloalkane dehalogenase LinB. *Chemical Research in Chinese Universities*, *25*(4), 556–559.

Zhou, X., Li, X., & Chen, X. (2013a). Binding mechanism of Orange G to human serum albumin: Saturation transfer difference-NMR, spectroscopic and computational techniques. *Dyes and Pigments*, *98*(2), 212–220. doi:10.1016/j.dyepig.2013.01.023

Zhou, X., Wang, Y., Hu, T., Or, P. M., Wong, J., Kwan, Y. W., ... Yeung, J. H. (2013b). Enzyme kinetic and molecular docking studies for the inhibitions of miltirone on major human cytochrome P450 isozymes. *Phytomedicine*, *20*(3), 367–374. doi:10.1016/j.phymed.2012.09.021 PMID:23102508

Zielinski, M., Kahl, S., Hecht, H. J., & Hofer, B. (2003). Pinpointing biphenyl dioxygenase residues that are crucial for substrate interaction. *Journal of Bacteriology*, *185*(23), 6976–6980. doi:10.1128/JB.185.23.6976-6980.2003 PMID:14617661

KEY TERMS AND DEFINITIONS

Ab initio: Method of estimation about biological features of a protein without any template structure or existing model.

Binding Mode: The positioning of ligand relative to the receptor and their coordination/conformation when bound/attached with each other.

De novo: Protein structure designed from the sequence alone.

Ligand: The molecule or entity which fits in/binds to the receptor. It might or might not be small in size compared to the receptor.

Pose: A specified position of ligand with respect to the receptor.

Ranking: Classification of poses based on scoring function.

Receptor: The molecule which acts as a receiving entity that might or might not be large in size as compared to the ligand.

Scoring: Evaluation of a pose through calculation of favorable interaction of parameters (such as hydrogen bonds, hydrophobic interactions).

This research was previously published in the Handbook of Research on Inventive Bioremediation Techniques edited by Jatindra Nath Bhakta; pages 24-45; copyright year 2017 by Engineering Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 79
The Performance of Educational Institutions Through the Electronic Records Management Systems:
Factors Influencing Electronic Records Management System Adoption

Muaadh Mukred
National University of Malaysia (UKM), Malaysia

Zawiyah M. Yusof
National University of Malaysia (UKM), Malaysia

ABSTRACT

This article discusses the relationship between educational institutes that use ERMS and the performance of those institutions. This article uses a mixed explanatory method that incorporated quantitative and qualitative approaches. The quantitative method collected the responses from 364 participants. This was followed by a qualitative approach where experts were interviewed to verify the model. The results generated using the quantitative approach demonstrated that the quality of the system, information, and service as well as the security provided by the system had a significant positive relationship with the successful adoption of ERMS, which in turn improved performance. Moreover, the qualitative results that gathered through the experts confirmed the findings and contributed to enriching the understanding of the adoption of ERMS in educational institutions.

DOI: 10.4018/978-1-7998-1204-3.ch079

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Records management strives to control the records and documents related to business processes, transactions and decision making (Mohd & Chell, 2005) as it manages records over their lifecycle (Yusof & Chell, 2000). There are several benefits to using a records management system. In general, a records management system provides reliable information that allows an organization to operate effectively (Asogwa, 2012; Mukred & Yusof, 2015) while encouraging transparency and accountability (Mohd & Chell, 2005; Yusof & Chell, 2000). In organizations that are strictly regulated, such as those in the private sector, an effective records management system may be required by law (Wright, 2013). In a related study, Egwunyenga (2009) stated that records are important to educational institutions as they help these organizations achieve their educational objectives, but the principles used to manage records in these environments are inadequate. Other studies confirm that there are issues associated with the accuracy, reliability, and trustworthiness of managing records. To be considered a "record" information must be valid and authentic (Mohammad & Yusof, 2013). It must also be understandable, obtainable, and secured. Records are used because they can facilitate planning and help make decisions. An effective records management system allows an organization to adhere to any applicable legal requirements, efficiently retrieve information, and it creates the infrastructure for a secure information system (Baron, Thurston, & mcleod, 2016).

Currently, in institutions of higher education, such as universities, there is a growing demand for environments that focus on governance, audits, and risk management. An Electronic Records Management System (ERMS) can meet the need for effective recordkeeping by ensuring that institutions of higher education will be efficient, maintain their records, remain trustworthy, encourage audits, become more competitive while meeting their fiduciary obligations and preserve their transparency and accountability (Atulomah, 2011). ERMS is described as the electronic management of e-records and non e-records with the help of computer equipment and software on the basis of acknowledged premises and practices of records management (Akor & Udensi, 2013).

To ensure that the adoption of an ERMS is successful, there are several factors that must be examined before an appropriate model can be developed (Mukred & Yusof, 2017). For example, in a developing nation such as Yemen, one of the challenges that must be overcome is inertia (Larson, 2009; Mukred, Yusof, Asma'Mokhtar, & Manap, 2016). The aim of this paper is to propose a model for the adoption of ERMS to support the organizations performance in Professional Education (HPE) institutions using the Information System Success Model by DeLone and McLean and Technology Acceptance Model (TAM). One of the primary goals of this study is to contribute to the discussion regarding the growth of e-records. This study also attempted to help institutions of higher education adopt ERMS. Additionally, this study also hopes to redress the lack of information in the literature regarding the relationship between ERMS and the performance of institutions of higher education.

ERMS AND EDUCATIONAL INSTITUTIONS

In order for an institution of higher learning to take advantage of the benefits associated with ERMS, staff must know how to use it properly. Unfortunately, as shown by a survey conducted by (Ismail, 2009), both the academic or the administrative staff at Kolej Yayasan Melaka in Malaysia lacked the skills and knowledge needed to access records management. Consequently, this college does not retain

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

their students' records. Records are difficult to manage if the organization does not have policies and frameworks in place to handle the management of their records. Mohammad and Yusof (2013) pointed out that policies and frameworks act as a roadmap when implementing records management procedures. These two researchers conducted a study in Melaka and Negeri Sembilan that examined vocational collage that had adopted a records management system without first developing relevant policies or frameworks. In a similar study, Galala and Yusof (2013) revealed that public post-secondary schools in Libya lacked policies regarding electronic records. Universities in Africa were found to have staff that were unaware of the organization's policies and procedures for using their ERMS (Bigirimana, Jagero, & Chizema, 2015).

Modern institutions of higher learning would benefit greatly from making governance, audits, and risk management a priority. This means that ERMS must ensure that these institutions can keep and maintain good records, and thus achieve their governance goals, while preserving their history, instilling trust, enhance their competitiveness, meet fiduciary requirements, and demonstrate transparency and accountability (Baron et al., 2016; Ghering, Caruso, & Gift, 2010). According to (Adu & Ngulube, 2016), effective recordkeeping is a fundamental component of good governance and more studies are required to better understand the role played by recordkeeping. They recommended that universities in South Africa and Malawi should consider the role of recordkeeping when they develop alternative frameworks for managing their records and documents. (Adu & Ngulube, 2016) highlighted that digital preservation must support organizations and meet any standards governing digital preservation.

The current literature reveals that most educational institutions in the developing world struggle to manage their records and documents even though they have improved their systems of governance. Educational institutions would benefit from change to policy, digital record storage, records management units, the support of senior management, education and training of records management staff, and systems and procedures for the retrieval and distribution of records (Phiri, 2016). ERMS are used in the developing word but still barriers to their successful implementation remain (Baron et al., 2016). The move towards using ERMS in education is worldwide because it minimizes the gap between the supply and demand for education. Several countries are aware that investing in the adoption of innovations, such as ERMS, can improve the whole educational system (Cheon, Lee, Crooks, & Song, 2012). ERMS contributes to the provision of education improving the performance of an educational institute.

PREVIOUS WORKS ON THE ADOPTION OF ELECTRONIC RECORDS MANAGEMENT SYSTEM

ERMS are relatively new tools for records officers and archivists in the developing world (Asogwa, 2012; Mosweu, Bwalya, & Mutshewa, 2017; Mukred et al., 2016). Currently, there are few studies that examine the role between information systems and decision making (Atulomah, 2011; Mukred & Yusof, 2017; Mukred et al., 2016). One reason for the lack of studies may be that there is a lack of tools for measuring how information systems affect decision making processes. Instead a standard benchmark is used for comparison. Qualitative and qualitative methods are needed to measure the quality of decisions. This requirement was the reason this study examined the impact of ERMS on the performance of educational organizations. In the following sections, earlier studies regarding the technical factors that could affect the implementation of ERMS and the role played by ERMS in organizations' performance are examined.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Studies that investigate ERMS in institutions of higher learning are limited but that are several studies that examine ERMS in healthcare. The records relied on by healthcare providers are managed by Electronic Health Records (EHR). The concepts, methods, and management techniques used by EHR are applicable to other sectors. The degree to which information technology is proportional to the degree to which the adopter is convinced that the new technology would lessen their workloads (Phillips, Calantone, & Lee, 1994). The adoption of technology has been described as the acceptance of technology that would appear to offer the best course of action (Rogers, 2010) and thus it focuses on the initial motivation and intentions surrounding the use of innovative technologies (Proctor et al., 2011). The adoption of ERMS improves educational delivery services. The process begins when a potential adopter becomes aware of a technology and it ends when the adopter implements and uses the technology (Renaud & Van Biljon, 2008).

System Quality

An information system is measured by the quality of the system's output (Delone & McLean, 2003). The quality is connected to its software and data components, and it assesses the health of the technical aspects of a system (Gorla, 2011) and it has a significant effect on its effectiveness, which can be defined as the ability of a system to do what it was designed to do. A study conducted by Zheng, Zhao, and Stylianou (2013) looked at the effect of system quality. In their study, they defined quality as the ability of a system to manage processes without crashing and it tested the readiness of a system to engage in its processes. In a recent study by Coffman, Spetz, Grumbach, Fix, and Bindman (2014), a system quality model was developed to determine the availability of e-health records (EHRs) by certified nurse midwives, nurse practitioners, and physicians in California. Their study took place after California Medicare implemented an EHR system. The researchers used similar features to gauge the availability of healthcare providers and they found that nurse midwives, practitioners and physicians were more likely to practice in areas with fewer Medicare patients.

Information Quality

A review of the literature found that there were several definitions for the quality of information. (Gorla, 2011) suggested that the quality of information is related to its perceived value in the eyes of the user. Gorla stated that information is perceived as being of high quality if it meets its specifications as well as the expectations and requirements of the user. (Pérez-Mira, 2010) divided information into two categories when they identified information as required or obtained. Required information meets the aesthetic, cognitive, functional, and technical requirements of users. (Pérez-Mira, 2010) speculated that the quality of information will significantly influence the effectiveness of an information system. Several studies reviewed in this paper demonstrated the significant influence information quality had on the system (Ali & Money, 2005; Mercy Mlay Komba & Ngulube, 2014; Mercy Milay Komba & Ngulube, 2015). Given its importance, quality of the information must be considered when selecting and adopting an information system (Petter, DeLone, & McLean, 2013). The quality of the information is related to the accuracy and knowledge of the system overall and any new information it might generate. One way to ensure that system requirements for information quality can be satisfied is to obtain feedback from end users (Ali & Money, 2005).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Service Quality

The quality of service is another variable examined in this study. Service quality refers to the difference between normal service expectations and the actual performance from perspective of the user. Service quality centers on the types of services provided (Gorla, 2011). Al-Hudhaif (2010) based service quality on the accuracy and reliability of a service. Reliability was demonstrated to be one of the most important characteristics of a quality service. In a digital world, reliability will improve the trustworthiness of an organization because it reflects the consistency and credibility of the organization (Hsiao, Li, Chen, & Ko, 2009). The quality of the service provided by an ERMS depends on its ability to meet the needs of educational institutions. Good quality service will encourage users to adopt and use a system (Motaghian, Hassanzadeh, & Moghadam, 2013).

Security

Safeguarding the security of electronic records is critical. Security can be better protected by administrators when they support security management and adopt initiatives (McGinn et al., 2012). A focus on security can lead to the development of robust, superior systems. The user's perception of the security of a system will determine if they will adopt a system. A study conducted by Chhanabhai and Holt (2007) revealed that users quickly accept ERMS systems when its security is guaranteed. The security risks associated with the adoption of a system must be taken into consideration, especially since the standards and legislative requirements related to the security of electronic records have increased (Daglish & Archer, 2009). Further research is needed to explore these regulations and the role they play in implementing secure systems (Fernández-Alemán, Señor, Lozoya, & Toval, 2013). Several studies have investigated the influence of security on the adoption of technology. In one study, Wright (2013) found that data protection should be standard. Public Key Infrastructure is a possible solution as it authenticates and encrypts data and can be very effective when coupled with sufficient memory and processing capacities.

Intention to Adopt ERMS

The intention to adopt refers to a user's determination to accept a new technology (Seymour, Makanya, & Berrangé, 2007). In this study, the intention to adopt is defined as the intention of a user to adopt ERMS and their attitude towards ERMS. Gong, Xu, and Yu (2004) pointed out that attitude alone is not enough to determine if a user will adopt a technology as they discovered that other factors, such as behavioral intention, played a greater role. In a similar study, the quality of the output of a system was more important than intention when predicting if physicians would adopt internet based health applications (Chismar & Wiley-Patton, 2003). In the developing world, the intention to adopt ERMS is influenced by the quality of the system, information, and service. If barriers to accepting and using ERMS are to be lifted, then a greater understanding of the factors that influence attitudes and ultimate acceptance of a system is required (AlAwadhi & Morris, 2008; Mukred et al., 2016).

Adoption of ERMS

A need exists to examine the technologies adoption that would assist educational institutions in transforming themselves into global leaders in HPE institutions. Understanding such an adoption is required

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to diffuse and urge the technology adoption as this may improve learning and teaching. Without such understanding, technologies adoption may fail, their effectiveness may become obsolete and their sustainability will become negligible (Nachmias & Ram, 2009), and in the current progressive and dynamic world, this could mean universities' lagging behind and the loss of valuable resources. Additionally, the ultimate aim of effective management for electronic records within HPE and universities requires that their staff accept and use ERMS and as such, this study examines the attributes/characteristics that impact academic members/staff of HPE institutions adoption of ERMS and that contribute to their organizations' performance. Delone and McLean (2003) examined how the quality of a system, its information and service influences a user's attitudes towards the system and their satisfaction. They developed a model that is now widely used to measure the success of the adoption of an information system. Their model has been independently tested (Tona, Carlsson, & Eom, 2012). This study used an updated version of their model to examine the influence of intention on the adoption of ERMS.

Organization Performance

In measuring performance, two measures are generally considered: primary measures, which is the degree to which an operation fulfils its objectives and secondary measures, which is the degree to which the needs of customers have been met (Slack, 2015). A critical effective management parameter is the measurement of performance measurement; without measuring something, it is difficult to improve it. Therefore, this paper aims to identify and measure the impact of ERMS practices on organizational performance in order to improve it. Good records management improves the performance of an organization as well as its competitiveness (Tagbotor, Adzido, & Agbanu, 2015). The improved performance and competitiveness of an organization can be measured by better decision-making process, increased productivity, more sales and rising profits, reduced costs, enhanced customer safety, as well as the creation of employment and economic development. According to (Urbach & Ahlemann, 2010), decision making satisfaction directly and positively affected the quality of the system and information in DeLone and McLean's (1992) model.

Chae, Yoo, Kim, and Chae (2011) studied the relationship between hospital information systems, Electronic Medical Records (EMR), Clinical Decision Support Systems (CDSS), and the performance of a hospital. They also examined how management issues changed over time and during different stages of growth. They used a chi-square test to investigate the characteristics of the EMR-CDSS relationship and a t-test to evaluate the effects of both systems on the performance of the hospital. (Chae et al., 2011) found that there was a significant relationship between EMR and CDSS thus indicating that both factors improved hospital performance. Moreover, (Plantier et al., 2017) investigated the relationship between electronic medical records (EMR) adoption and organizational performance in French hospitals. The study concluded that EMR adoption directly affected the performance of the organization. The findings suggest that there should be an investment in EHRs for better organizational performances.

It is evident from the literature that studies addressing the factors influencing ERMS adoption in support of the organizations' performance (Smith, Bradley, Bichescu, & Tremblay, 2013); (Osakwe, 2013) are still few and far between, with an evident lack of framework that could guide in ERMS proper adoption, and in the system's role in the performance of HPE institutions. This calls for an urgent need to examine the organizations' performance through ERMS adoption. Exploration and investigation need to be carried out in order to put ERMS in place in educational setting to help in its decision-making process and organizations' performance.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DEVELOPING ERMS ADOPTION MODEL

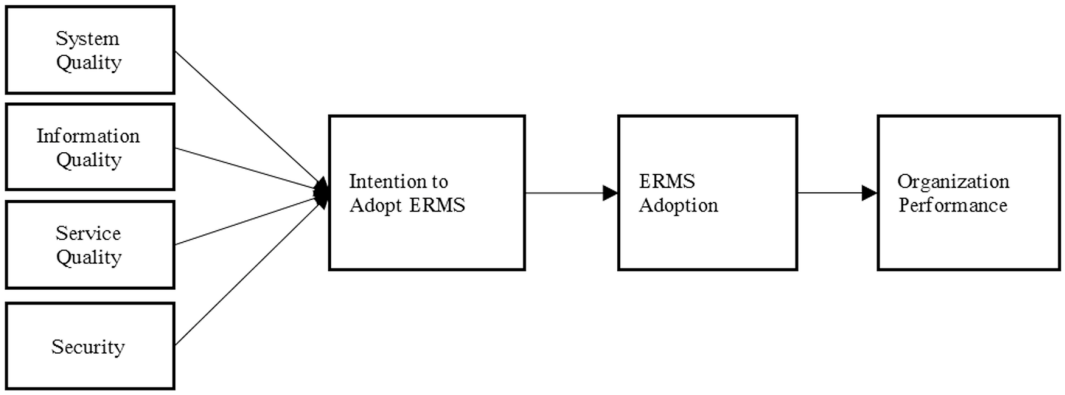
If universities are to become global educational leaders, they must identify those technologies that would help them meet their goals. Without this information, there is a risk that they may fail to adopt the right technology and become ineffective before becoming unsustainable and obsolete (Nachmias & Ram, 2009). The effective management of records in an educational institute means that staff must not only accept ERMS, but they must also use the system. Consequently, in this study, the attributes and characteristics of the staff and educators in institutes of higher learning and their effect on the adoption of ERMS was examined.

The literature points out the inability of institutions of higher learning to adopt ERMS. One of the primary goals of this study was to address the issues surrounding ERMS adoption and to develop an ERMS model that would facilitate organizations' performance. The model developed by this study was the result of analyzing the technical factors that impacted the adoption of the system by staff. DeLone and McLean's (2003) model was employed to evaluate the staff's intention to adopt ERMS and its relationship with organizations' performance. In other studies that looked at the adoption of ERMS, TAM, with or without modifications, was used to analyzed the adoption or acceptance of a ERMS (Gagnon et al., 2010; Gagnon et al., 2014; Morton & Wiedenbeck, 2010; Archer & Cocosila, 2011; Marcu, 2013; Terry, Brown, Denomme, Thind, & Stewart, 2012). Based on the success of previous studies, this study also used the TAM (1993) model to gauge how well the ERMS was adopted and its effect on the performance of an organization. Combining TAM with DeLone and McLean's (2003) model has been recommended (Cheok & Wong, 2015) to investigate organizational performance in an educational environment. The proposed model is illustrated in Figure 1.

The following hypothesis were tested in this study:

H1: The quality of the system positively influences the intention to adopt the ERMS.
H2: The quality of the information positively influences the intention to adopt ERMS.
H3: The quality of the service positively influences the intention to adopt ERMS.
H4: The security of the system positively influences the intention to adopt ERMS.
H5: The intention to adopt has a positive relationship with ERMS Adoption.
H6: The adoption of ERMS positively affects the performance of an organization.

Figure 1. The proposed model for ERMS adoption



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

METHODOLOGY

This study used an organized set of logical phases. Choosing between a quantitative, qualitative, or a combined approach depends on the research objectives (Creswell, 2013). Quantitative research approaches are best used with objective theories that can explore the relationships between variables. By contrast, qualitative research approaches are appropriate when the research objectives are concerned with individuals, groups of individuals, as well as social or human issues. In many studies related to information systems, neither a quantitative or qualitative approach, but combined approaches, are becoming more common because they can provide valuable information (Ngulube & Ngulube, 2015). Creswell (2013) described three different types of combined research methods. These methods are the explanatory, exploratory, and triangulation methods. The explanatory method conducts an initial quantitative study and then it uses a follow-up qualitative study. The exploratory mixed-method uses a quantitative study that relies on the results from an earlier qualitative study. The triangulation method, which is the most common approach, interprets data gathered using both qualitative and quantitative methods. The goal of the triangulation method is to confirm the results of a study using both quantitative and qualitative methods (Creswell, 2013).

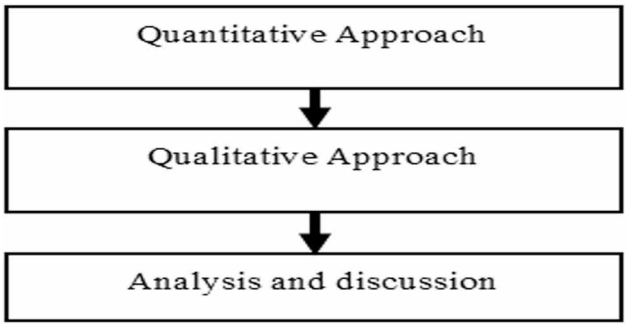
Harris and Brown (2010) noticed that in studies that used mixed methods, the structured questionnaires and semi-structured interviews produced confirmatory results. Likewise, (Ngulube & Ngulube, 2015) observed that mixed method allowed researchers to paint a comprehensive picture of a phenomenon and realize their research objectives. The clear picture and valuable insight provided by the data from studies that employ a mixed methodology allow for the development of more accurate theories (Ngulube & Ngulube, 2015). This study used an explanatory mixed method to collect data. First, quantitative data was gathered from 364 respondents working in either managerial or academic positions at the Ministry of Technical Education and Vocational Training (MTEVT) in Yemen. Myers (1997) asserted that in the social sciences, qualitative research techniques allow researchers to examine social and cultural phenomena. He stated that qualitative methods further our understanding of individuals and the social and cultural aspects of their environment. Qualitative studies are important because they provide insight into an individual's perceptions and behaviors through observations and questions (Glaser & Strauss, 2009). Collecting and analyzing qualitative data provides information about a phenomenon. In this study, qualitative data was used to better understand the role played by ERMS adoption in the performance of an organization. The most commonly used qualitative data collection methods are interviews, documentation, and observations (Myers, 1997). Interviews are typically used to gather data in qualitative studies and they may be combined with quantitative data collection methods to support the results. In this study, qualitative data was collected from the participants using interview methods.

The interview method is used to collect the experiences and perceptions of small group of interviewees (Brenner, Green, & Camilli, 2006). Typically, and interviewer asks the participants open ended questions about the topic being studied. The answers provided by the interviewees reflect their personal experiences and they are recorded (Creswell, 2013). The results of the interviews provided valuable information that is very different than the static data collected using quantitative methods. Additionally, interviews ensure the consistency of the quantitative results in studies that employ mixed methodologies. The explanatory approach starting with quantitative methods then qualitative methods is shown in Figure 2.

Figure 2 illustrates that the first step in the Explanatory Mixed Method is to take a quantitative approach. The data from the survey was used to observe the relationships between exogenous and endogenous constructs (Kline, 2015), predict and control phenomena, and determine the best concept (McNabb, 2015).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Explanatory mixed method



The data gathered using a quantitative approach is analyzed using mathematical procedures or statistics (Creswell, 2013). Qualitative data collection techniques should be related to the research objectives. In this study, this step included designing a questionnaire. Questionnaires are a common tool used by researchers thus it is important that any questionnaire used is effective. (Sekaran, 2006) recommended that particular attention should be paid to the wording of the questionnaire, how items are categorized, the scales and codes used to analyze the responses to the questionnaire, and the general appearance of the questionnaire. In this study, care was taken that the questionnaire protected the quality of the data, received enough responses, and that any anticipated biases were mitigated (Groves et al., 2011).

This study uses the online questionnaire format where emails were sent to clarify the purpose behind the research and to invite the participants to complete it. (Peytchev, Couper, McCabe, & Crawford, 2006) stated that online format surveys have several benefits over paper-based methods. For instance, an online survey takes two-thirds of the time taken to complete a traditional paper-based survey (Groves et al., 2011). The researcher does not have to wait to retrieve the paper questionnaires as data is automatically gathered. Online questionnaires also decrease the costs used and due to the automatic processing or responses, the results can be accessed in a timely manner (Schonlau, Ronald Jr, & Elliott, 2002). The questionnaire is divided into five sections with the contents as follows; the first section includes questions about the measurement of the information quality, system quality, service quality and security factors are contained; in the second section, the measurement of the behavioral intention factor are contained. The third section contained the questions regarding the adoption of ERMS and in the fourth section, the measurement of performance factors are contained. The respondents were requested to indicate their level of agreement to the items concerning the adoption of ERMS through a five-point Likert scale that ranged from strongly disagree depicted by 1 to strongly agree depicted by 5. Finally, the fifth question contains the participants' general information in terms of their demographic characteristics – these are entered into descriptive statistics to provide an insight into the study sample.

Studies that use mixed methodologies generally employ purposive sampling to collect qualitative data (Palinkas et al., 2015). In purposive sampling, units or cases are chosen for a reason rather than being selected at random. For example, participants may be selected for their knowledge or other characteristics. Identifying participants can ultimately help researchers better understand a phenomenon (Creswell, 2013). The standard for purposive sampling is that participants should be knowledgeable about the subject. Current studies of ERMS adoption that focus on developing countries highlight the importance of interviewing individuals from different levels of management because the definition of "success" depends on the subjective perspective of each individual (Heeks, 2006). An interview should begin by introduc-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ing the topic being examined as well as the purpose and scope of the study. In this study, the questions asked during the interviews were related to validating the ERMS model. The data from the interviews was used to validate the model. Five experts were questioned during a semi-structured interview. The interviewer asked open-ended questions so that the experts could freely express their opinions. The interviewer also asked questions related to the factors that effected the successful adoption of ERMS. All the questions in the interviews helped the researcher comprehend the perceptions of the participants regarding the role of ERMS in organizations performance. The goals of the interviews were as follows:

1. Confirm that the factors highlighted in the literature were valid for ERMS and that the terminology could be translated into Arabic;
2. Determine if the study model was researchable;
3. Determine the importance of relevant factors;
4. Determine if other factors described by the practitioners were omitted by earlier studies;
5. Develop items from the design of the questionnaire.

Each interview session was conducted in a period lasting 35-55 minutes. The interview sessions were recorded, after which it was transcribed, and the transcripts were verified by a second person to determine inconsistencies. The highlighted minor inconsistencies were rectified by modifying the transcripts.

RESULTS AND DISCUSSION

Before the data can be analyzed, the constructs must be validated. Factor analysis is used to confirm exogenous, endogenous, and dependent variables. In addition to factor analysis, this study also confirmed the reliability of the items used to measure the constructs. An exploratory factor analysis (EFA) was undertaken using principal component analysis (PCA) with an orthogonal VARIMAX rotation to examine how well exogenous items completed the following five tasks:

1. The first task was to determine the number of factors that affected the exogenous, endogenous and dependent variables. If the factor had an eigenvalue greater than 1.0, it indicated that the factor had a measurable influence on a variable. By contrast, if the factor had an eigenvalue less than 1.0, it meant that that factor did not have a significant impact (Hair Jr, Hult, Ringle, & Sarstedt, 2016);
2. The second task used a Kaiser-Meyer-Olkin (KMO) value to measure sampling adequacy. The adequacy of a sample is confirmed if it had a value that exceeded 60% (Pallant, 2013);
3. In the third task, the correlation and partial correlation values were used to confirm that the size of each sample was adequate. A correlation value greater than 50% indicated that the size of the sample was adequate (Field, 2009);
4. The fourth task required that the significance communality, which indicated the significance of an item, was determined. Only factors with a communality percentage greater than 50% were considered significant (Tabachnick & Fidell, 2013).

In this study, structural equation modeling (SEM) was used to validate the model. The path coefficients were calculated using Smart PLS 2.0. Smart PLS 2.0 was also used to determine the significance of the path using a bootstrapping method. As recommended by (Hair Jr et al., 2016), a composite reli-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ability (CR) value greater than 0.7 indicates internal consistency and convergent validity (measured using Average Variance Extracted (AVE)) must be 0.5 or greater. Furthermore, the Cronbach's alpha value must be greater than 0.7 and factor loading should result in a value of at least 0.6. Table 1 sets out the results of the SEM analysis.

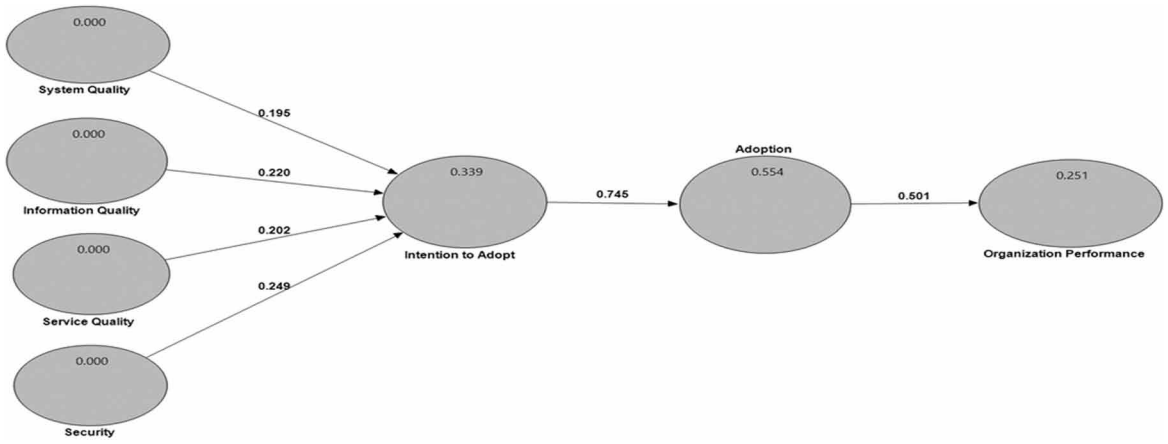
Table 1 shows that the values were sufficient (Hair Jr et al., 2016). The values for factor loading ranged from 0.748 to 0.928 and the Cronbach's alpha ranged from 0.828 to 0.906. The CR values ranged from 0.869 to 0.930 indicating that the model had good internal consistency. The AVE values were between 0. 571 and 0. 741. The path model is illustrated in Figure 3.

This study employed a DeLone and McLean IS model and TAME model and the results described the relationship between the exogenous, endogenous, and dependent variables. The results for each research hypotheses are summarized in Table 2.

As shown in Table 2, the quality of the system, information, and service and security determined the likelihood that an ERMS would be adopted. Because ERMS plays an important role in the decision-making processes of an educational institution and its performance, these organizations should ensure that the ERMS they adopt have the features (system, information, and service quality and security) that will encourage their employees to embrace its use. The findings indicated that the technical factors should be taken into consideration as a determinant of ERMS adoption in HPE sector and that it influences individuals and organization alike in different levels (operational, tactical, and strategic). Respondents stressed the crucial aspects of the system characteristics, including system quality, information quality and service quality in the system adoption as well as security. In addition, findings confirmed the positive impact of ERMS on organizations' performance. Once the statistical model was validated using SEM, the proposed model was further validated by five IS experts to confirm that the model would achieve its purpose. The five experts had experience in both academic, practical, and government settings. These experts validated the completeness of the model, the relationships between variables, correlations, and the strength of the relationships found in the proposed model. They also provided an explanation of the adoption model in institutions of higher learning.

The potential for bias originating from a single discipline was eliminated by selecting interview participants with different disciplinary backgrounds. As a result, the participants in the interviews included application integration managers, planning and coordination department managers, and business analysts.

Figure 3. Path model for the model proposed in this study



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Results of PLS analysis

Construct	Item	Loading	Cronbach's Alpha	Composite Reliability	AVE
System Quality	SYSQ1	0.748	0.879	0.912	0.675
	SYSQ2	0.783			
	SYSQ3	0.848			
	SYSQ4	0.858			
	SYSQ5	0.863			
Information Quality	INFQ1	0.812	0.862	0.900	0.643
	INFQ2	0.814			
	INFQ3	0.816			
	INFQ4	0.785			
	INFQ5	0.780			
Service Quality	SRVQ1	0.829	0.828	0.896	0.741
	SRVQ2	0.888			
	SRVQ3	0.866			
Security	SQRTY1	0.737	0.843	0.887	0.612
	SQRTY2	0.804			
	SQRTY3	0.810			
	SQRTY4	0.789			
	SQRTY5	0.769			
Intention to Adopt	IADOP1	0.827	0.867	0.904	0.654
	IADOP2	0.843			
	IADOP3	0.713			
	IADOP4	0.822			
	IADOP5	0.834			
Adoption	ADOPT1	0.728	0.828	0.869	0.571
	ADOPT2	0.748			
	ADOPT3	0.764			
	ADOPT4	0.768			
	ADOPT5	0.769			
Organization's Performance	ORGPR1	0.928	0.906	0.930	0.728
	ORGPR2	0.812			
	ORGPR3	0.805			
	ORGPR4	0.832			
	ORGPR5	0.884			

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Summary of the results for the research hypotheses

Hypothesis		ß	Support	R^2
Intention to Adopt				0.339
H1	System Quality →Intention to Adopt ERMS	0.195	Supported	
H2	Information Quality →Intention to Adopt ERMS	0.225	Supported	
H3	Service Quality →Intention to Adopt ERMS	0.202	Supported	
H4	Security →Intention to Adopt ERMS	0.245	Supported	
Adoption				0.554
H5	Intention to Adopt → ERMS Adoption	0.745	Supported	
Organization's Performance				0.251
H6	ERMS Adoption → Organization's Performance	0.501	Supported	

The clarity of translated materials was confirmed by discussing the results in both Arabic and English with a professor with experience in IS. The results of the interviews revealed that the quality of the ERMS system was a reflection of the quality and accuracy of the information as described by Expert 1:

The most attractive advantage that promotes the ERMS use is that it provides quality information that satisfies the need of the decision makers in a timely manner. Your model is interesting owing to the clear and sufficient output information from the ERMS that is good evidence that information quality is a reflection of system quality.

Expert 2 described just how crucial quality information is to ERMS adoption:

The information quality factor in the model is crucial due to the dependence of decision making on the quality of available information.

Access to information technology requires appropriate infrastructure. If IT solutions such as ERMS are to work, access to them must be easy and reliable. The interview participants saw service quality as an important factor for adopting ERMS as expressed by Expert 3:

Service quality factor is important to enable employees and administrators to adopt the system. Your proposed model defines the system applicability even for future implementation.

The benefits of using ERMS include access to the records needed to make informed decisions as stated by Expert 4:

Information extracted from e-records through ERMS is vital for effective decision making.

Expert 3 made a similar statement:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

For easy decision-making, information has to be made available and this is made possible through extraction of information using ERMS.

As did Expert 5:

Information availability and system accessibility influence the ERMS adoption and implementing the system in HPE institution is important as this could play a significant role in evaluating OBE programs, and in enhancing the process of decision making. In other words, the ERMS enables time and cost saving.

The importance of ERMS to the managerial and academic activities of an educational institute was highlighted by Expert 1 who stated:

Electronic records management system is important for successful evaluation process in the outcome-based education. The system's adoption will help in saving cost, time and effort that will improve organizations' performance, from the academic viewpoint. From the viewpoint of administration, the system enables administrative processes depending on the system's generation of relevant information and will impact the organizations performance as well.

The findings were also supported by the experts, who confirmed that system factors (system, information, and service quality and security) play key roles in the adoption of ERMS in their HEP institutions. Similar result was reported by (Al Haderi, 2014); (Tzeng, Kuo, Talley, Chen, & Wang, 2015); (Mercy Milay Komba & Ngulube, 2015); (Sharma, 2015); (Ghazvini & Shukur, 2013); Compeau and Terry (2017). However, what makes this study unique is the proposal of an integrated model for those factors and their impact on the intention to adopt ERMS and the relationship between such adoption and organizations performance. The validation process included obtaining feedback from the experts through interviews (qualitative method) to support the results. The findings validated ERMS model as reliable and applicable in HPE institutions. The validity of the proposed model and the support on all the hypotheses are largely attributed to the involvement of experts.

CONCLUSION

One of the goals of this study was to identify the factors that would positively influence the adoption of ERMS in institutions of higher learning. This goal was met by developing the model shown in Figure 3. This model was assessed using SEM and Smart PLS statistical software. The results indicated that the model was valid. The model used in this study represents a theoretical breakthrough because it included variables associated with adopting innovations. An effective ERMS results in information that can help educational institutions make well informed decisions and improve their performance in a competitive marketplace. The results of this study confirm that ERMS significantly improves the performance of organizations. The completeness of the ERMS model used in this study was confirmed by a panel of experts because the model summarized the importance of those factors that increased the likelihood that an ERMS would be successfully adopted to the benefit of the organization. Furthermore, this study narrowed the gap in the literature related to the relationship between ERMS adoption and organizational performance by using both the DeLone and McLean IS Success Model and TAM model. The results

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

from this study also contribute to the literature regarding the validation of the DeLone and McLean and TAM model for institutions of higher learning in Yemen and other developing countries. Consequently, this study could shape future theory and empirical studies concerning ERMS, especially in cases where it is used to improve organizational performance. In addition, further research is called to build on the findings of this study, particularly to identify the other factors such as organizational and technological factors that might influence the adoption of ERMS.

ACKNOWLEDGMENT

We would like to thank Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia by giving the authors an opportunity to conduct this research. This research is funded by Universiti Kebangsaan Malaysia under Research Grant Code: GUP-2017-046

REFERENCES

Adu, K. K., & Ngulube, P. (2016). Key threats and challenges to the preservation of digital records of public institutions in Ghana. *Information Communication and Society*, 1–1.

Akor, P. U., & Udensi, J. (2013). An assessment of record management system in establishment division of two universities in Nigeria. *International Letters of Social and Humanistic Sciences*, *13*, 97–109. doi:10.18052/www.scipress.com/ILSHS.13.97

Al Haderi, S. M. (2014). System characteristic facilitates the acceptance of information technology in middle east culture. *International Journal of Business and Social Science*, *5*(6).

Al-Hudhaif, S. A. (2010). Measuring quality of Information System Services in manufacturing organizations in Riyadh. *Economics and Administration, 24*(1).

AlAwadhi, S., & Morris, A. (2008). The Use of the UTAUT Model in the Adoption of E-government Services in Kuwait. *Paper presented at the 41st Annual Hawaii International Conference on System Sciences*. 10.1109/HICSS.2008.452

Ali, A. S. B., & Money, W. H. (2005). A study of project management system acceptance. *Paper presented at the 38th Annual Hawaii International Conference on System Sciences*. 10.1109/HICSS.2005.53

Archer, N., & Cocosila, M. (2011). A comparison of physician pre-adoption and adoption views on electronic health records in Canadian medical practices. *Journal of Medical Internet Research*, *13*(3), e57. doi:10.2196/jmir.1726 PMID:21840835

Asogwa, B. E. (2012). The challenge of managing electronic records in developing countries: Implications for records managers in sub Saharan Africa. *Records Management Journal*, *22*(3), 198–211. doi:10.1108/09565691211283156

Atulomah, B. C. (2011). Perceived records management practice and decision making among university administrators in Nigeria.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Baron, J. R., & Thurston, A. (2016). What lessons can be learned from the US Archivist's digital mandate for 2019 and is there potential for applying them in lower resource countries? *Records Management Journal*, *26*(2), 206–217. doi:10.1108/RMJ-04-2015-0015

Bigirimana, S., Jagero, N., & Chizema, P. (2015). *An Assessment of the Effectiveness of Electronic Records Management at Africa University*. Zimbabwe: Mutare.

Brenner, M. E., Green, J., & Camilli, G. (2006). Interviewing in educational research. In Handbook of complementary methods in education research (Vol. 2).

Chae, Y. M., Yoo, K. B., Kim, E. S., & Chae, H. (2011). The adoption of electronic medical records and decision support systems in Korea. *Healthcare Informatics Research*, *17*(3), 172–177. doi:10.4258/hir.2011.17.3.172 PMID:22084812

Cheok, M. L., & Wong, S. L. (2015). Predictors of e-learning satisfaction in teaching and learning for school teachers: A literature review. *International Journal of Instruction*, *8*(1), 75–90. doi:10.12973/iji.2015.816a

Cheon, J., Lee, S., Crooks, S. M., & Song, J. (2012). An investigation of mobile learning readiness in higher education based on the theory of planned behavior. *Computers & Education*, *59*(3), 1054–1064. doi:10.1016/j.compedu.2012.04.015

Chhanabhai, P., & Holt, A. (2007). Consumers are ready to accept the transition to online and electronic records if they can be assured of the security measures. *Medscape General Medicine*, *9*(1), 8. PMID:17435617

Chismar, W. G., & Wiley-Patton, S. (2003). Does the extended technology acceptance model apply to physicians. *Paper presented at the 36th Annual Hawaii International Conference on System Sciences*. 10.1109/HICSS.2003.1174354

Coffman, J. M., Spetz, J., Grumbach, K., Fix, M., & Bindman, A. B. (2014). Electronic health record availability among advanced practice registered nurses and physicians. *The American journal of managed care, 20*(11), eSP31-38.

Compeau, D. R., & Terry, A. (2017). Connecting medical records: An evaluation of benefits and challenges for primary care practices. *Journal of Innovation in Health Informatics*, *24*(2), 204–215. doi:10.14236/jhi.v24i2.855 PMID:28749315

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Daglish, D., & Archer, N. (2009). Electronic personal health record systems: a brief review of privacy, security, and architectural issues. *Paper presented at the World Congress on Privacy, Security, Trust and the Management of e-Business CONGRESS'09*. 10.1109/CONGRESS.2009.14

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30. doi:10.1080/07421222.2003.11045748

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Egwunyenga, E. J. (2009). Record Keeping in Universities: Associated Problems and Management Options in South West Geo-Political Zone of Nigeria. *International Journal of Educational Sciences*, *1*(2), 109–113. doi:10.1080/09751122.2009.11889983

Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. Á. O., & Toval, A. (2013). Security and privacy in electronic health records: A systematic literature review. *Journal of Biomedical Informatics*, *46*(3), 541–562. doi:10.1016/j.jbi.2012.12.003 PMID:23305810

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Gagnon, M.-P., Ghandour, E. K., Talla, P. K., Simonyan, D., Godin, G., Labrecque, M., ... Rousseau, M. (2014). Electronic health record acceptance by physicians: Testing an integrated theoretical model. *Journal of Biomedical Informatics*, *48*, 17–27. doi:10.1016/j.jbi.2013.10.010 PMID:24184678

Gagnon, M.-P., Ouimet, M., Godin, G., Rousseau, M., Labrecque, M., Leduc, Y., & Abdeljelil, A. B. (2010). Multi-level analysis of electronic health record adoption by health care professionals: A study protocol. *Implementation Science; IS*, *5*(1), 30. doi:10.1186/1748-5908-5-30 PMID:20416054

Galala, K. A., & Yusof, Z. M. (2013). Electronic Records Management in Institutions of Higher Learning in Libya: A Case Study. *Journal of Information & Knowledge Management*, *12*(01), 1350001. doi:10.1142/S0219649213500019

Ghazvini, A., & Shukur, Z. (2013). Security challenges and success factors of electronic healthcare system. *Procedia Technology*, *11*, 212–219. doi:10.1016/j.protcy.2013.12.183

Ghering, C., Caruso, J. B., & Gift, D. (2010). Electronic Record Management: Today's High Stakes. *ECAR Research Bulletin*, *8*, 1–12.

Glaser, B. G., & Strauss, A. L. (2009). *The discovery of grounded theory: Strategies for qualitative research*. Transaction publishers.

Gong, M., Xu, Y., & Yu, Y. (2004). An enhanced technology acceptance model for web-based learning. *Journal of Information Systems Education*, *15*(4), 365.

Gorla, N. (2011). An assessment of information systems service quality using SERVQUAL+. *ACM SIGMIS Database*, *42*(3), 46–70. doi:10.1145/2038056.2038060

Groves, R. M., Fowler, F. J. Jr, Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011).. . *Survey Methodology*, *561*, ●●●.

Hair, J. F. Jr, Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.

Harris, L. R., & Brown, G. T. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data.

Heeks, R. (2006). Health information systems: Failure, success and improvisation. *International Journal of Medical Informatics*, *75*(2), 125–137. doi:10.1016/j.ijmedinf.2005.07.024 PMID:16112893

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hsiao, S.-J., Li, Y.-C., Chen, Y.-L., & Ko, H.-C. (2009). Critical factors for the adoption of mobile nursing information systems in Taiwan: The nursing department administrators' perspective. *Journal of Medical Systems*, *33*(5), 369–377. doi:10.100710916-008-9199-8 PMID:19827263

Ismail, N. (2009). *Kajilan mengenai pekiskisan program pergurusan rekod di Kolej Yayasan Melaka*. UKM.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Komba, M. M., & Ngulube, P. (2014). An Empirical Application of the DeLone and McLean Model to Examine Factors for E-Government Adoption in the Selected Districts of Tanzania. *Emerging Issues and Prospects in African E-Government*, 118.

Komba, M. M., & Ngulube, P. (2015). *Factors That Influence E-Government Adoption in Selected Districts of Tanzania*.

Larson, S. (2009). Information Systems and strategic decisions: A literature. *RE:view*.

Marcu, A. D. (2013). Relationship of Self-Efficacy to the Stages of Concern in the Adoption of an Innovation in Higher Education.

McGinn, C. A., Gagnon, M.-P., Shaw, N., Sicotte, C., Mathieu, L., Leduc, Y., ... Légaré, F. (2012). Users' perspectives of key factors to implementing electronic health records in Canada: A Delphi study. *BMC Medical Informatics and Decision Making*, *12*(1), 1. doi:10.1186/1472-6947-12-105 PMID:22967231

McNabb, D. E. (2015). *Research methods for political science: Quantitative and qualitative methods*. Routledge.

Mohammad, N., & Yusof, Z. M. (2013). Nurturing knowledge management initiative base on students' electronic records: A study in vocational colleges in Melaka and Negeri Sembilan zones. *Paper presented at the 2013 International Conference on Research and Innovation in Information Systems (ICRIIS)*. 10.1109/ICRIIS.2013.6716715

Mohd, Z., & Chell, R. W. (2005). *Issues in records management*. Penerbit Universiti Kebangsaan Malaysia.

Morton, M. E., & Wiedenbeck, S. (2010). *EHR acceptance factors in ambulatory care: a survey of physician perceptions. EHR Acceptance Factors in Ambulatory Care: A Survey of Physician Perceptions/ AHIMA*. American Health Information Management Association.

Mosweu, O., Bwalya, K. J., & Mutshewa, A. (2017). A probe into the factors for adoption and usage of electronic document and records management systems in the Botswana context. *Information Development*, *33*(1), 97–110. doi:10.1177/0266666916640593

Motaghian, H., Hassanzadeh, A., & Moghadam, D. K. (2013). Factors affecting university instructors' adoption of web-based learning systems: Case study of Iran. *Computers & Education*, *61*, 158–167. doi:10.1016/j.compedu.2012.09.016

Mukred, M., & Yusof, Z. M. (2015). The Role of Electronic Records Management (ERM) for supporting Decision making Process in Yemeni Higher Professional Education (HPE): A Preliminary Review. *Jurnal Teknologi*, *73*(2), 117–122. doi:10.11113/jt.v73.4202

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mukred, M., & Yusof, Z. M. (2017). The DeLone–McLean Information System Success Model for Electronic Records Management System Adoption in Higher Professional Education Institutions of Yemen. *Paper presented at the International Conference of Reliable Information and Communication Technology*.

Mukred, M., Yusof, Z. M., Asma'Mokhtar, U., & Manap, N. A. (2016). Electronic Records Management System Adoption Readiness Framework for Higher Professional Education Institutions in Yemen. *International Journal on Advanced Science*, *Engineering and Information Technology*, *6*(6), 804–811.

Myers, M. D. (1997). Qualitative research in information systems. *Management Information Systems Quarterly*, *21*(2), 241–242. doi:10.2307/249422

Nachmias, R., & Ram, J. (2009). Research insights from a decade of campus-wide implementation of web-supported academic instruction at Tel Aviv University. *The International Review of Research in Open and Distributed Learning*, *10*(2). doi:10.19173/irrodl.v10i2.595

Ngulube, P., & Ngulube, B. (2015). Mixed methods research in The South African Journal of Economic and Management Sciences: An investigation of trends in the literature. *South African Journal of Economic and Management Sciences*, *18*(1). doi:10.17159/2222-3436/2015/v18n1a1

Osakwe, R. (2013). The Efficiency of Managing School Records by Secondary School Principals in Delta State, Nigeria. *Annals of Modern Education*, *5*(1), 72–80.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, *42*(5), 533–544. doi:10.100710488-013-0528-y PMID:24193818

Pallant, J. (2013). *SPSS survival manual*. UK: McGraw-Hill Education.

Pérez-Mira, B. (2010). Validity of Delone and Mclean's model of information systems success at the web site level of analysis. Northwestern State University.

Petter, S., DeLone, W., & McLean, E. R. (2013). Information systems success: The quest for the independent variables. *Journal of Management Information Systems*, *29*(4), 7–62. doi:10.2753/MIS0742-1222290401

Peytchev, A., Couper, M. P., McCabe, S. E., & Crawford, S. D. (2006). Web survey design: Paging versus scrolling. *International Journal of Public Opinion Quarterly*, *70*(4), 596–607. doi:10.1093/poq/nfl028

Phillips, L. A., Calantone, R., & Lee, M.-T. (1994). International technology adoption: Behavior structure, demand certainty and culture. *Journal of Business and Industrial Marketing*, *9*(2), 16–28. doi:10.1108/08858629410059762

Phiri, M. J. (2016). *Managing university records and documents in the world of governance, audit and risk: Case studies from South Africa and Malawi*. University of Glasgow.

Plantier, M., Havet, N., Durand, T., Caquot, N., Amaz, C., Philip, I., ... Perrier, L. (2017). Does adoption of electronic health records improve organizational performances of hospital surgical units? Results from the French e-SI (PREPS-SIPS) study. *International Journal of Medical Informatics*, *98*, 47–55. doi:10.1016/j.ijmedinf.2016.12.002 PMID:28034412

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., ... Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health*, *38*(2), 65–76. doi:10.100710488-010-0319-7 PMID:20957426

Renaud, K., & Van Biljon, J. (2008). Predicting technology acceptance and adoption by the elderly: a qualitative study. *Paper presented at the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology*. 10.1145/1456659.1456684

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Schonlau, M., Ronald, D. Jr, & Elliott, M. N. (2002). *Conducting research surveys via e-mail and the web*. Rand Corporation.

Sekaran, U. (2006). *Research methods for business: A skill building approach*. John Wiley & Sons.

Seymour, L., Makanya, W., & Berrangé, S. (2007). End-users' acceptance of enterprise resource planning systems: An investigation of antecedents. *Paper presented at the 6th Annual ISOnEworld Conference*.

Sharma, S. K. (2015). Adoption of e-government services: The role of service quality dimensions and demographic variables. *Transforming Government: People*. *Process and Policy*, *9*(2), 207–222.

Slack, N. (2015). *Operations strategy*. Wiley Online Library. doi:10.1002/9781118785317.weom100160

Smith, A. L., Bradley, R. V., Bichescu, B. C., & Tremblay, M. C. (2013). IT governance characteristics, electronic medical records sophistication, and financial performance in US hospitals: An empirical investigation. *Decision Sciences*, *44*(3), 483–516. doi:10.1111/deci.12019

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics.

Tagbotor, D. P., Adzido, R. Y. N., & Agbanu, P. G. (2015). Analysis of Records Management and Organizational Performance. *International Journal of Academic Research in Accounting*. *Finance and Management Sciences*, *5*(2), 1–16.

Terry, A. L., Brown, J. B., Denomme, L. B., Thind, A., & Stewart, M. (2012). Perspectives on electronic medical record implementation after two years of use in primary health care practice. *Journal of the American Board of Family Medicine*, *25*(4), 522–527. doi:10.3122/jabfm.2012.04.110089 PMID:22773720

Tona, O., Carlsson, S., & Eom, S. (2012). An Empirical Test of DeLone and McLean's Information System Success Model in a Public Organization.

Tzeng, W.-S., Kuo, K.-M., Talley, P. C., Chen, H.-C., & Wang, J.-J. (2015). Do ePortfolios Contribute to Learners' Reflective Thinking Activities?: A Preliminary Study of Nursing Staff Users. *Journal of Medical Systems*, *39*(9), 100. doi:10.100710916-015-0281-8 PMID:26254803

Urbach, N., & Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *Journal of Information Technology Theory and Application*, *11*(2), 5.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Wright, T. (2013). Information culture in a government organization: Examining records management training and self-perceived competencies in compliance with a records management program. *Records Management Journal*, *23*(1), 14–36. doi:10.1108/09565691311325004

Yusof, Z. M., & Chell, R. W. (2000). The Records Life Cycle: An inadequate concept for technology-generated records. *Information Development*, *16*(3), 135–141. doi:10.1177/0266666004240413

Zheng, Y., Zhao, K., & Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision Support Systems*, *56*, 513–524. doi:10.1016/j.dss.2012.11.008

This research was previously published in the International Journal of Information Technology Project Management (IJITPM), 9(3); edited by John Wang; pages 34-51; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 80
Online Social Exergames for Seniors:
A Pillar of Gamification for Clinical Practice

Stathis Th. Konstantinidis
NORUT, Norway

Josef Hallberg
Luleå University of Technology, Sweden

Ellen Brox
NORUT, Norway

Gunn Evertsen
NORUT, Norway

Per Egil Kummervold
NORUT, Norway

Johannes Hirche
Luleå University of Technology, Sweden

ABSTRACT

The population is getting older, and the resources for care will be even more limited in the future than they are now. There is thus an aim for the society that the seniors can manage themselves as long as possible, while at the same time keeping a high quality of life. Physical activity is important to stay fit, and social contact is important for the quality of life. The aim of this chapter is to provide a state-of-the-art of online social exergames for seniors, providing glimpses of senior users' opinions and games limitations. The importance of the motivational techniques is emphasized, as well as the impact that the exergames have to seniors. It contributes to the book objectives focusing on current state and practice in health games for physical training and rehabilitation and the use of gamification, exploring future opportunities and uses of gamification in eHealth and discussing the respective challenges and limitations.

DOI: 10.4018/978-1-7998-1204-3.ch080

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

An increase in the life expectancy and decrease in birth rate cause an ageing population and according to the UN reports the rate of population ageing in the 21st century will exceed that of the previous century. Older people (aged 60 years or over) increased from 9.2 per cent in 1990 to 11.7 per cent in 2013 and will continue to grow as a proportion of the world population, reaching 21.1 per cent by 2050. A continuous increase on the number of "oldest old" (aged 80 year or over) is expected, reaching 19 per cent in 2050 from 14 per cent in 2013 (United Nations - Department of Economic and Social Affairs Population Division, 2013).

Healthcare is costly and all countries face budget constraints, so spending more money is not the solution. Policy makers and health professionals believe that the solution lies in an active and healthy ageing (*White Paper: "Together for Health: A Strategic Approach for the EU 2008-2013"*, 2007). Services for each individual's whole-person need are emerging and should prolong the time elderly individuals can live autonomously and healthily (Berry & Bendapudi, 2007).

The use of information and communication technology (ICT) is an emerging field in active and healthy aging trying to create better conditions of life for the older adults and to strengthen the industrial opportunities. This includes the use of games for keeping seniors active and healthy in the form of exergames and gamified applications in combination with sensors and motion devices.

This new era of research gives rise to a multitude of challenges. Techniques, methods and strategies that are already applied in different domains are gathered in order to provide state-of-the-art games which will enhance the physical activity of seniors. Despite the fact that a lot of studies engage research on how a health game contribute to behaviour change, patients' self-management of care, and adherence to care, remain an open research field (Baranowski, Buday, Thompson, & Baranowski, 2008; Honka, Kaipainen, Hietala, & Saranummi, 2011; D. King, Greaves, Exeter, & Darzi, 2013; Strong, Tulu, Agu, & He, 2014) in order to be applied properly in the target group of seniors. Persuasive games and motivational strategies are starting to be applied in games for physical activity for seniors (Ellen Brox, Luque, Evertsen, & Hernandez, 2011; Kort, IJsselsteijn, Eggen, & Hoven, 2005), but complete frameworks which prove their effectiveness are still missing.

This Chapter is structured as follows. Initially an overview over games that are used by seniors for exercise are described. Mainly proprietary but also a few open-source games can be categorized in game types, providing the current trends. Many of those are normally not played online, although some offer online solutions too. Gamification is also being researched as a current trend in exergaming. The following section will identify the existing research that takes (or took) place through relevant EU and national projects. A review of those projects reveals the research gaps which need to be fulfilled in the future. An analysis of general requirements and an investigation into the different technology possibilities is following. The next section is focusing on the impact of exergames for seniors. This section strengthens the knowledge regarding improvements that exergames can have for seniors, thus providing the ground to build on. Physical activity can also improve the mental health and will be included in the key issues for possible positive contribution of exergames. The following section will focus on social aspects which have been proved to be an important factor of everyday life, especially for seniors. This section records the existing efforts on the inclusion of this aspect into the exergames. Motivation, on the other hand, is really important to keep the seniors to frequently play the exergames. Persuasive techniques that are used in exergaming are recorded and the connection with social aspects is emphasized. Last but not least future challenges and research questions will be presented in order to further boost the rapidly developing area of games in healthcare.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

TRENDS OF EXERGAMES FOR SENIORS

Inactivity is the most common reason for loss of functional capabilities among elderly. Physical activity can effectively delay this process, and is one of the most important factors in promoting wellbeing among elderly (Goldspink, 2005; Taylor et al., 2004). However, adherence to exercise programs is a major obstacle. It requires large human resources to have long term effects (Teixeira, Carraca, Markland, Silva, & Ryan, 2012). It is in this light exergames are promising.

Exergames can be defined as "digital gaming systems with an interface that requires physical exertion to play" (Larsen, Schou, Lund, & Langberg, 2013). With exergames, one is thinking of digital devices that both record and instruct the users how to exercise correctly and motivate the users by giving feedback on their achievements to encourage them.

There are several types of exergames on the marked but they are not targeting seniors. Still some senior centres and nursing homes around the world have exergames as an offer for their users. It seems that mainly the two major proprietary game consoles are being used; namely Nintendo Wii and Kinect for Xbox by Microsoft. All home entertainment systems that have a way of registering users' physical activities can be used for exergames.

There is a multitude of games for all of these platforms, but in general they are targeted towards younger users. Some of the games are however also suitable for older users and are offered in the senior centres that use exergames. However speed, information overload, language, etc. make most of these games challenging for the seniors.

The Nintendo Wii mainly has two different controls used for elderly; a handheld remote or a balance board that register change in weight distribution. The remote is normally held in one hand, and several of the games can also be played while seated. There are also walking / jogging games that are played where the remote is supposed to be in a pocket, and it is possible to buy an extra remote ("nunchuck") to have one remote in each hand (used in for instance cycling games). Particularly bowling with the Wii remote has been considered suitable for elderly since it is turn based and one can play in one's own speed. But also several of the balance board games are suitable for seniors, and can be played with a chair as support and even a walking frame if the balance board fits underneath it. Touchdown's Dancetown Fitness System® was a dance mat for seniors with a frame to hold on to, but that has been discontinued.

The Kinect camera for Xbox offers a hands free gaming experience since there are no remotes or physical devices to be handled during gameplay. This could therefore be very suitable for seniors, and some of the games can also be played by elderly. However many current Kinect games are too fast or otherwise not suitable (for instance they require jumping).

Studies show that elderly seem to enjoy playing commercial Wii games (Chao, Scherer, Wu, Lucke, & Montgomery, 2013; Williams, Soiza, Jenkinson, & Stewart, 2010). Several studies also show that gameplay can be good for balance and rehabilitation (Billis et al., 2010; Chao et al., 2013; Jung, Li, Gladys, & Lee, 2009; Rendon et al., 2012; Theng, Dahlan, Akmal, & Myint, 2009; Williams et al., 2010; Wollersheim & Merkes, 2010). Another important aspect is socialising. Earlier studies show a larger psychological and physical well-being of seniors playing Nintendo Wii games compared to traditional board games due to the type of intervention with peers (Jung et al., 2009). There are however not many exergames that are designed for this age group, and particularly not commercially available. It is worth to note that the motions and interactions are all made for a healthy (and younger) audience (Anderson,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Annett, & Bischof, 2010; Chao et al., 2013), scores and progress measurements are too generic for rehabilitation use (Anderson et al., 2010), many of the games are difficult to play (Anderson et al., 2010) and many games are too fast (Theng et al., 2009; Wollersheim & Merkes, 2010).

Since Wii was on the marked before Kinect, there is much more literature about seniors and Wii play. However, since developing for Wii is hard, there are only a few games created for elderly on Wii like the Fit4All (Billis et al., 2010; E. I. Konstantinidis, Billis, Hlauschek, Panek, & Bamidis, 2010). In 2012 a study on motion sensors applied to exergames for elderly (Hors Fraile, Browne, Brox, & Evertsen, 2012) concluded that Kinect was a better option than Wii. Commercial Kinect is now being used for regular training for seniors, and studies on the use of Kinect for elderly have also started to emerge (Maggiorini, Ripamonti, & Zanon, 2012; Miyosawa, Yanagisawa, & Okuhara, 2013).

Another trend is to use accelerometers to register movements, for instance step-counters. There is however concern about the usefulness of this type of equipment for some of the elderly. Many elderly have gaits with hardly any leg movements, so step counters will not register movements correctly. Also wristbands that are growing in popularity are challenging for those using walking frames (Lauritzen, Muñoz, Luis Sevillano, & Civit, 2013).

Gamification

According to Wikipedia, Gamification is the use of game thinking and game mechanics in non-game contexts to engage users in solving problems. This is also used for physical exercise. McCAllum (McCallum, 2012) states that serious games and gamification refer to the way the experience has been designed. If it is an addition of game elements to an existing activity, it is gamification, but if it requires creating a whole new experience it is a game. Since most people walk, adding game elements to walking, for instance using a step counter can be considered gamification. There are many apps made for step counters, both dedicated step counters and phone apps, and many apps offer sharing of your effort online, for instance on Facebook. These can be considered gamification of walking, since it adds elements of competition and rewards (for instance likes of peers) to activities people already do. This also utilises peer pressure, since you share your walking – and scores – with others.

Gotsis, M. et al (Gotsis, Wang, Spruijt-Metz, Jordan-Marsh, & Valente, 2013) have tried out a web-based physical activity diary for adults, also with social gaming features, and they have positive results. When the diary adds feedback and rewards (even just stars on the screen) this can be considered gamification. (Stinson et al., 2013) designed and developed a smartphone-based pain diary app that would be engaging to adolescents with cancer through the gamification of pain assessment recordings. The gamification process involved adolescents playing the role of law-enforcement officers on a special investigative force that hunts down pain. Another example of a gamification can be found in the "Lifestyle Tool" intervention (Thorsteinsen, Vittersø, & Svendsen, 2014). A pilot test of an interactive, computer-tailored website physical activity intervention, with additional SMS text messages has gamified components: participants received points for completed activities; the test designed to increase the motivation for being physically active; and helping participants to set goals: social contract and competition. This study suggests that using gamification in physical activity interventions is a fun and engaging way to help people achieve behaviour change. The effect seems to drop off after some weeks though, but further research is needed (Thorsteinsen et al., 2014).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In addition gamification, applied with mHealth (mobile health) applications, is an encouraging implementation for incentivizing improved patient self-management according to (Miller, Cafazzo, & Seto, 2014). Some preliminary studies as reported by (Hamari, Koivisto, & Sarsa, 2014) indicate that gamification provides positive effects, but these effects are greatly dependent on both the context in which the gamification is being implemented and on the users using it.

Walking is an activity that elderly do, but the apps that exist are not elderly friendly. Within the GameUp project, to motivate the seniors, a gamification app was made; the users defined both individual and group goals, and there were game rewards when goals were reached. The step counter (Fitbit) also registered height differences, so one of the group goals was to climb a nearby mountain. Groups in two countries (Norway and Spain) could compare results. Both the common goals and the comparison with another group turned out to be motivational. Also an exerbiking game can be considered gamification for those who use it with stationary bikes they already have.

Gamification can be used to motivate elderly to continue to do what they do, or to do more of what they are able to do. Walking, even indoors, is one obvious activity, but also different kinds of housework can be beneficial. A roleplaying game called HabitRBG (habitrpg.com/) – gamify your life – offers users to define daily task, habits (good and bad) as well as to-do tasks. This game, or one using the same concept, can also be used for exercising.

Online Social Exergames for Seniors

Online social exergames for seniors could be defined as exergames that are played together with people that are not in the same room. Typically this requires that the participants are online at the same time, but there are also examples of turn-based games or games with other social elements (scorebords, common goals, etc) that do not have this requirement. Gamification could also be considered as another type of social online interaction. For instance gaining rewards for physical activities like going for a walk, vacuum-cleaning, etc.

Games can also be divided into collaborative and competitive games. In collaborating gameplay, two or more users must collaborate to solve a task. Multiplayer games are also often divided into Real Time Games and Turn Based Games (Schaefer, Enderes, Ritter, & Zitterbart, 2002). Typically the participants can do multiple moves/actions in Real Time Games and the game ends in a set time or when a particular goal is reached, for instance Wii tennis. In Turn Based Games, participants alternate in making moves. The advantage of turn based games is that all players do not have to be present at the same time, for instance Wii or Kinect Slalom. Wii Tennis on the other hand is an example of a turn-based game where both players have to be present simultaneusly.

Many games use mixed methods, for instance some turn based games are also supposed to be played together. Bowling is one such example, where players take turns, but they wait for each other and watch the other play.

The main purpose of motion games is the motions of the player – that is to make the players physically active. The elderly players get very involved in the exergames, but many commercial games can be a challenge for several reasons (Anderson et al., 2010; Theng et al., 2009; Wollersheim & Merkes, 2010). Some of these are speed, too much information at the time since elderly only manage to focus on what to do (correct movements to play the game and not additional information), colours and contrasts, etc. These points are especially important when dealing with elderly persons. Since the player has problems

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

focusing on things outside the gameplay (Ellen Brox et al., 2011; Gerling, Schild, & Masuch, 2010), it poses some challenges when integrating social elements. Online social exercise games for elderly therefore need very simple game interfaces.

The most common devices that are used for motion capturing or monitoring are shown in Table 1. Two of them are tracking the entire body, while the majority are using sensors to capture the movement of the parts of the body where the sensor is placed. New products coming mainly for smart TVs are able to identify the movement of the hand and it is possible to use it to perform some excersises. Each of the products have pros and cons and usually the choice is made according to the exergame or the gamified application.

Glimpses of Older Adults' Reaction to Online Social Exergames

Four senior colleagues (55+) at Norut were recruited to try a collaborative commercial Kinect game online called "1000 leaks". Only one out of the four had previously tried Kinect games. The trial was conducted in two rooms, and the chat function was on. The main aim of the trial was to see how the

Table 1. Existing products for motion tracking

Product	Body Part Tracking	Known Issues
PlayStation Eye camera	All the body	Development stop
Microsoft Kinect/ASUS/PrimeSense cameras	All the body	• Bad in daylight • Leg in camera
Nintendo Wii/Wii mote hand held remote	hand movement legs movement (depend on sensor placement)	
Wii Fit Balance Board	weight distribution	Does not track feet placement
Dance pads	feet placement	
PlayStation® Move hand hold remote	hand movement	
Open source Smartphones	Accelerometer (depend on sensor placement)	Short battery life
Android Smartphones	Accelerometer (depend on sensor placement)	Short battery life
Apple IOS Smartphones	Accelerometer (depend on sensor placement)	Short battery life
Apple iPhone 5	compass, accelerometer and gyroscope (depend on sensor placement)	• Tracks when phone is asleep, • Allows pedometer apps to run in the background, • Share results with friends nearby
Wearable sensors in health and fitness	(depend on sensor placement)	
Leap Motion's 3D gesture control	only hand and fingers	
Extreme Motion software with standard 2D camera	upper part of body	Good in near mode, good in daylight
PointGrab's software with standard 2D camera	hands	Connected with Smart TV platforms
Future Smart TV with cameras and gesture-recognition software	hands	Forthcoming technology

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

online gaming was received. Some early results showed that all found it fun, but those who had not played before found it a bit strange at first, and it was difficult to chat or listen to the other online players. It was also not so easy to understand that the other player was a real person in the next room. But all three who had not played before would like to play more. Despite the fact that the sample is really small, the results provide indicators for future research areas. This small trial empasizes the issue of the importance of integration of the social elements outside the gameplay, and encourages the users to talk, chat, or send messages/images between game sessions and also compare results.

A walking app based on Fitbit™ was developed in the GameUp project. This app both gave the users the possibility to define personal goals as well as offering groups of elderly to define group goals. This was tested in a senior centre where the elderly would have the fitbit at home for two weeks. It turned out that this was very motivational for the elderly, and some who could not go for a walk outdoors would rather walk much more inside their homes.

Effects and Barriers of Exergames for Seniors

Exergames for elderly have been the topic for multiple research projects. However, most of the projects have been isolated attempts at improving the situation, and most of them have been experiments with very few participants. The last couple of years, there have however been a couple of serious attempts at compiling the evidences of the effect of exergames in review articles.

The (Bleakley et al., 2013) review study is aiming to research the effect of computer games for elderly. The review looks at more than a thousand articles about "Interactive Computer Games" (ICG) for older adults (aged over 65) and excludes those that are related to rehabilitation after an injury. It ends up with a scoping review of twelve studies. The study concludes that there is "*preliminary evidence that ICG is a safe and effective medium to promote physical activity in an older population, and may be associated with a range of physical and cognitive benefits"*.It also concludes that, "*there was no evidence of an optimal exercise dose in terms of exercise type, duration, or intensity and, there were insufficient data to investigate potentially important subgroup effects based on age, gender, or baseline activity level"*. The authors in another study (Larsen et al., 2013) started to screen 1861 articles and ended up with 7 that fulfill their criteria which were randomized controlled trials about exergames for "healthy elderly". They concluded that all the included studies have low to medium methodological quality. Six out of the seven studies had positive results for the intervention. They are however too heterogeneous to make comparisons possible.

The typical trial size in exergame trials is between 20 and 40 participants. Several of the studies were able to prove that training is better than no training. The studied effect size is large enough to detect a difference with this sample sizes. However, very few of the studies where able to prove that exergames are better than traditional training. This might be caused by a lack of effect or that the studies are underpowered. It should also be noted that none of the studies had done pre-trial power tests. These tests are used to determine whether you have enough participants to detect a difference if such difference exists. Furthermore, there are no indications that exergames are worse than other training programs. In many of the studies there are also possible effects of exergames both in cost reductions and accessibility. Instead of studying the clinical effects, it is possible to aim the study at these aspects.

An important barrier to all these advancements could be the digital literacy and more specifically the gaming skills of seniors. It is common knowledge that the digital divide between seniors and younger adults is too wide (Olphert & Damodaran, 2013). When it comes to health-related ICT use (Wiklund

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Axelsson, Melander Wikman, Näslund, & Nyberg, 2013) health digital literacy is even lower. Nonprofit organisations (e.g. SeniorNet - http://www.seniornet.org/), municipalities, Universities of the Third Age, community information centers, ICT associations for seniors, and home teaching usually take the role of a tutor and provide basic computer-using knowledge to the seniors (Nguyen, Tapanainen, & Obi, 2014). To the best of our knowledge there are no researches showing who should teach gaming skills to seniors. But, based on the authors' experience and projects, gaming skills could be taught in senior centres, by visiting volunteers, by formal or informal carers or even by children or grandchildren. It could also be done in rehabilitation clinics if for instance exergames is being used there as part of the therapy. How it is done in each case will depend on the state of the person and how the elderly care is organised in that country or region.

Seniors with (mild) cognitive impairment are an important piece of the pie of potential users of the online exergames. The cognitive impairment of this group could act as a barrier to access the online exergames. However, even cognitively healthy elderly persons can have problems with technology. A person who cannot longer remember how to perform tasks she has done most of her life, will not be able to operate the equipment required for exergames. People with dementia go through many phases, and can master many things in the beginning, but nonetheless this target group will need help and support. However, physical and social activities are also important for people with cognitive problems. Simple games with activities, music and images from "back then" can maybe be used to activate persons in this group.

Game design targeting cognitive impaired seniors should consider the two major approaches which have been used successfully to train older adults directly in cognitive abilities:

1. Strategy training; and
2. Extended practice training.

For instance digital action games use general principles similar to those of extended practice training (Zelinski & Reyes, 2009). In addition, there is preliminary evidence that interactive computer games is a safe and effective exercise intervention for an older population, and may be associated with a range of physical and cognitive benefits (Bleakley et al., 2013; Zelinski & Reyes, 2009).

CURRENT AND PAST PROJECTS IN EXERGAMING FOR ELDERLY

The European Union is actively engaged in promoting active and healthy ageing (European Union, 2010) through law, initiatives and research funding. The Commission has proposed a Decision (1312/2013/EU, 2013) on the Strategic Innovation Agenda of the European Institute of Innovation and Technology (EIT) for 2014-2020 identifying "Innovation for healthy living and active ageing" as one of the priority themes for the EIT Knowledge and Innovation Communities (KICs) wave in 2014-2015. To this extent research grants are being continuously announced in the form of call for proposals towards active and healthy ageing. AMBIENT ASSISTED LIVING JOINT PROGRAMME (http://www.aal-europe.eu) is a funding activity that aims to create better condition of life for the older adults and to strengthen the industrial opportunities through the use of information and communication technology (ICT).

Similarly, strategies for active aging are presented and the needs for healthy and independent seniors are also defined in the rest of the developed countries (British Columbia. Ministry of Health, 2005). For example in the United States of America a report from the department of Health identifies the need

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to pursue health-promoting behaviours in older people and to meet the growing needs and demands of a rapidly aging nation (Centers for Disease Control and Prevention, 2013). In Australia development of strong public policy is the key issue to influence healthy ageing (Renehan et al., 2012). Also in the South-East Asia Region the (*Yogyakarta Declaration on Ageing and Health 2012*, 2012) sets out a number of policy commitments to be implemented in each member state. The Association of Southeast Asian Nations (ASEAN) Strategic Framework for Social Welfare and Development (2011 – 2015) (The ASEAN Secretariat, 2012) encompasses a number of initiatives to improve the lives of older people. The ASEAN has also recognised that the rapidly ageing population presents challenges to society to adapt social and economic policies.

There are numerous projects focusing on physical training and mobility of seniors, but less on the use of exergames, serious games and gamified applications in order to provide a motivational environment for senior users to keep exercising. There are several related projects since physical activity for the aging population of Europe is a theme that is receiving great attention due to its importance both for the target group and our societies. Table 2 is depicting EU co-funded projects targeting the improvement of physical activity of seniors by the use of exergames or gamified application and comparing them according to the type of game they are offering:exergames or games that have a game engine and that are controled with motion capture camera or other motion sensors; also whether the games are online and social. Whether they provide an administration portal for the carers of the elderly is also included.

Many of the projects in Table 2 are in addition using sensors such as acelometers, custom movement systems, heart rates, etc. The ELF@HOME system will comprise the wearable physical activity sensor(developed by EFL@HOME), biomedical sensors, a simple TV interface for fitness sessions and a computer vision system to analyse fitness exercises execution. All these components are connected to a service platform implementing the intelligence needed (Carús, García, García, Waterworth, & Erdt, 2014). To this extent, the PAMAP system aims to support out-of-hospital physical rehabilitation and physical activity monitoring services. The PAMAP system consists of the sensor network (accelerometers and heart rate monitor sensor) and a control unit that links the sensors and several different I/O devices. An interactive TV (i-TV) interface is used to enable patients to gain access to their health record and the details of the rehabilitation plan to be followed (Reiss, Stricker, & Lamprinos, 2012; Reiss & Stricker, 2011). Another project, FoSIBLE, was aiming to improve mental and physical well-being as one of its many goals by providing physical exercises based on different tracking and acceleration sensor technologies (Alaoui, Ting, & Lewkowicz, 2014).

Some gamified applications do not use any kind of sensors, but they create apps that encourage seniors to perform activity and provide useful information. The project DOSSy is targeting outdoor activities by developing an app for mobile devices. A hiking guide, high quality route information and basic safety system, which constantly checks the surrounding conditions (weather, daylight) for critical situations has been created ("DOSSy:Digital Outdoor and Safety System [project website]," 2012). The project MOBILE.OLD belongs to the same category. This project aims to provide a combined smartphone and TV-based service infrastructure with residential and outdoor services that will be delivered in a highly personalised and intuitive way and will advance the mobility of older persons (Karlhuber & Schürz, 2013). The MOB MOTIVATOR project aims to provide a comprehensive and engaging environment for promoting mobility and cognitive skills amongst elderly people, in both an indoor and an outdoor multiplayer gaming environment, which enables collaboration in an all-age social network ("momo: Mobility Motivator [project webpage]," 2012). Within the Is-Active project a gamebased app that make

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. EU co-funded projects aiming at the improvement of physical activity of seniors by the use of exergames or gamified application

Project Acronym: Title/Web Address	Exergame	Gamified Application	Online Game	Administration Portal
GameUp: Game-based mobility training and motivation of senior citizens www.gameupproject.eu	X	X		X
Nettrim: Social online exercises for elderly nettrim.norut.no	X		X	
Exergaming for balance training of elderly citizen www.imdi-sprint.nl/research/huidige-projecten/exergaming-project/?lang=en	X			X
ELF@Home:Elderly sELF-care based on sELF-check of health conditions and sELF-fitness at home elfathome.eu		X		X
IS-ACTIVE: Inertial Sensing Systems for Advanced Chronic Condition Monitoring and Risk Prevention www.is-active.eu		X		
PAMAP: Physical Activities are essential for Rehabilitation and Well-Aging www.pamap.org		X		X
RGS: The Rehabilitation Gaming System rgs-project.eu	X			
iStoppFalls: ICT-based Systems to Predict & Prevent Falls www.istoppfalls.eu	X			X
FARSEEING: FAll Repository for the design of Smart and sElf-adaptive Environments prolonging INdependent livinG farseeingresearch.eu	X			
Join IN: Senior Citizen Overcoming Barriers by Joining Fun Activities join-in-for-all.eu/	X	X		X
LLM: Long Lasting Memories www.longlastingmemories.eu	X	X		X
SILVER GAME http://www.aal-europe.eu/projects/silver-game/	X			X
MOB MOTIVATOR or Momo: Motivational health aid www.mobility-motivator.uvsq.fr	X		X	
Fit-TV fit-tv.nl		X		
DOSSy Digital Outdoor and Safety System dossy-aal.iwi.unisg.ch		X		
MOBILE.OLD: Residential & outdoor services advancing the mobility of older persons www.mobiledotold.eu		X		X
FoSIBLE: Fostering Social Interaction for a Better Life of the Elderly fosible.eu/		X		
REWIRE: Rehabilitative Wayout in Responsive home Environments www.rewire-project.eu	X			X

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

use of an accelometer was developed. During the gamification of walking by using an app (people agree to walk at the same time, but not physically together), feedback is given about the score, pulse rate, and oxygen saturation for motivation and a safe exercising (Tabak & Marin-Perianu, R.S. Hermens, 2012).

The latest projects engage the use of games that have a game engine and are controlled with motion capture camera or motion controls. The iStoppFalls project developed an embedded system that can predict and prevent falls by monitoring mobility-related activities and other risk factors of falls in real-life. Beyond continuous fall risk monitoring, this enables tailoring individualized exercise programs coached by iStoppFalls. In addition, an Xbox/Kinect based fall preventive exergame facilitates real preventive exercise training and biomechanical modelling and optional heart rate data are used (Meurer & Wieching, 2012; Wieching et al., 2012). Within the FARSEEING project an exergame using motion capture camera was developed for seniors (Nawaz et al., 2014). The game has nature as a theme, and it is targeting to improve strength and balance and reduce the risk os falls. The Silvergame Project (according to the AALwebsite http://www.aal-europe.eu/projects/silver-game/) develops a platform for serious gaming to foster the social inclusion of elderly. It aims at making a prototype that will include three different modules on one central platform: a silver song club where people can sing together, a multimedia driving simulator for cognitive training and a dance and fitness application. In addition, Fit-TV in the Netherlands is using a Wii remote around a foot and films shown on a TV screen to let people cycle through a virtual landscape. They can use a stationary bike or a set of pedals placed in front of the chair of the user (Otten, 2012). The Long Lasting Memories (LLM) project developed a platform which comprised of three existing interoperable components that perform complementary and interactive tasks to provide the system's services: the Physical Training Component (PTC); the Cognitive Training Component (CTC) and the Independent Living Component (ILC). The PTC component is based on the Fit For All system developed by the Lab of Medical Informatics of the Aristotle University of Thessaloniki. Fit For All is a game platform that can help elderly people to exercise and maintain their physical status and well-being through exergames, using Wii Balance Board and Wii controls (P. D. Bamidis et al., 2011; P. Bamidis et al., 2011; E. I. Konstantinidis et al., 2010). Another example of an early prototype that uses Wii controls for exergames is the Join-In project. In this project the main aim was that the users should be able to join fun activities online, including exergames. A cycling app and early prototypes of 3D walking and exercising games for Kinect (E. Brox, Evertsen, Horst-Fraile, Browne, & Åsheim-Olsen, 2015) have been developed (Cantwell, Broin, Palmer, & Doyle, 2012; Johnsen, Burkow, & Vognild, 2012).

A few of the projects combine both gamified applications and exergames. For example the GameUp project developed a walking app to improve the endurance of the seniors who use a commercial step counter. Both individual and group common goals are provided to the seniors and insertion of the data into the app allow them to see their performance and whether the individual and team group goals are reached or not. Also, a physical exergame has been developed in order to improve balance, range of movement and strength. Capture of the user movements is done with a motion capture camera. To this extent, the Nettrim project has been analysing and further expanded one of the games developed in GameUp with social elements. Since Nettrim is a pre-project it aims to establish the basis for international research projects targeting elderly and physical activity (S. T. Konstantinidis, Fernández-Luque, Brox, & Kummervold, 2014).

The Rehabilitation Gaming System (RGS) developed a Virtual Reality (VR) tool for the rehabilitation of deficits that occur after brain lesions and has been successfully used for the rehabilitation of the upper extremities after stroke. The RGS uses a motion sensor camera to track body movements and Data

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

gloves to track finger movements and control games on a screen. The project does not refer directly to physical activity for all the seniors, but only for people after stroke, but it represents a good example of virtual reality tools used in physical activity (Cameirão, Badia, Oller, & Verschure, 2010). In the category of projects which are not only targeting seniors, but patients with stroke, the REWIRE project provides a set of mini-games shown on the TV screen, controlled by a motion capture camera, a balance board and a haptic device. Social aspects were implemented, providing communication through web 2.0 tools. (Massé, Paraschiv-Ionescu, Ženko, Džeroski, & Aminian, 2013; Wüest, Langenberg, & Bruin, 2013)

There are not many projects that have an online aspect of interaction between users within the game. Some of them are focusing on social interaction, but mainly through traditional communication technologies. The role of formal and informal carers is emphasized by half of the projects. The importance of monitoring and redefinition of the performed exercises is fostered through administration portals, where the carer could intervene and adjust the tension of the excersise and in some case also the type of the excersises according to the senior's personal needs.

Furthermore there are a few games that are targeting seniors. Some of them are already out on the market like the Dancetown Fitness System (http://www.popcitymedia.com/innovationnews/dancetown1212.aspx) which is a slow motion version of "Dance Dance Revolution". The game system comes with a frame to prevent the players from falling and provide the opportunity to level up into a steady paced dancing for those mastering the "oldies" level. Others are experimental platforms that are available for use, such as the KinectLabs (http://myweb.polyu.edu.hk/~kinelabs/) which offer coordinated upper limb, lower limb and trunk balance training tasks and also support multi-player mode, but they are still in prototype phase. On the other hand a few proprietary products have appeared such as FysioGaming (http://doctorkinetic.nl/en/fysiogaming/) but they are also in the first stage, without proper validation of their impact.

There are also research efforts for exergames or gamified application which include a robot or a robotic tool in order to motivate the seniors to perform their physical training (Dubowsky, Genot, Godding, Kozono, & Skwersky, 2000; Heller, Senior, & Wheat, 2014; Wada, Shibata, Saito, & Tanie, 2004). However, those exergames or gamified applications usually require more infrastructure than a motion detection sensor and they are mainly targeting rehbilitation of seniors (Novak, Nagle, & Riener, 2014; Tamura et al., 2004) rather than keeping the seniors active and healthy.

IMPACT OF EXERGAMES FOR ELDERLY

If everyone started to eat healthily, gained control over their weight, started to sleep a sufficient amount, managed their stress, stopped smoking, and limited their alcohol intake to a reasonable level, then 90% of all type-2 diabetes, 80% of all cardio-vascular diseases, and 70% of stroke cases would be prevented (Honka et al., 2011). It has also been shown that lack of physical and social activity, stress, poor diet, and poor sleep quality contribute along-side of the genes, to the onset and worsening of dementia (Hallberg, J., Casati, 2013). Many of the more common chronic diseases, such as cardio-vascular disease, type-2 diabetes, obesity, and dementia, are therefore dependant on staying active, both physically and socially. In some cases a healthy life-style change can completely relieve the symptoms of the disease, and perhaps even cure it. In other cases it is a matter of being able to live with the disease in such a way that the disease does not get worse, and so that the life-quality is maintained. Exergames for Elderly provide the necessary platform for people to meet and share a good and social time together, and to help motivate each other to stay physically active.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Physical activity is also good for the mental health (Rosenberg et al., 2010), and can help relieve stress. However it is also plain to see that physical exercise will be good for our bodies more generally (not just our brains) and as such can help to combat the many diseases that flourish because of our modern high-caloric sedentary life styles. We now have to worry about stress related diseases, cardiovascular disease, diabetes, metabolic syndrome and so on—diseases that were almost unheard of in earlier societies. Exercise does not only have a direct and even immediate positive impact on brain health, but it also gives us a release from stress and gives us a way to expend that enormous surplus of calories that many of us ingest on a daily basis.

Stress, worry and depression are other problems many elderly face. The problem of prolonged stress has become recognized within the medical community. One consequence of stress can even be physically damaged brains. Stress can literally kill neurons in extreme cases. But even if this does not happen, there is evidence of other sorts of damages. Stress can lead to compromised neurons in the hippocampus. That is they do not die, but their dendritic trees shrivel up. If the person suffers then the brain eventually also is likely to suffer.

Here we also couple our social life to stress. We are, at a very fundamental level, social beings. We not only have a need to be social, we are social primates as a fact of biology. Not surprisingly, it is becoming clear that having an active social life is a key factor for brain health and cognitive endurance. As other primates, we reduce anxiety and stress in social settings that are safe to us. Research indicates that those who maintain active social lives tend to be less likely to develop dementia. Many dangerous health conditions develop because of prolonged stress. It is when we are unable to turn of stress responses that our health suffers.

"Almost all stressed out people state social anxiety and relationship difficult as their primary causes of stress. Broadly, social anxiety comes from three major relationship categories – child-parent, marital, and career/professional" ("Stress and Relationships," 2013).

As stated in the previous section both physical and social activity can help relieve stress. These activities can therefore also contribute to a happier life. The relationship between happiness and health is also perceived by people. For example, a qualitative study reported in (Kaplan, Madden, Mijanovich, & Purcaro, 2013) showed that participants saw a strong causal relation between stress and unhealthy behaviors, including in particular eating and sleeping habits as well as smoking. On the contrary, happier individuals tend to take life-preserving measures. A study of over 300,000 individuals in the US has shown that happier people are significantly more likely to wear seatbelts, and that the effect persists even when taking into account many other variables such as gender, marital status, income, or exercise level (Goudie, Mukherjee Sr., De Neve, Oswald, & Wu, 2011).

Diener and Chan (Diener & Chan, 2011) provide a comprehensive overview of dozens of studies on the topic. What seems to emerge is the presence of a causal relation between positive affect and a reduced occurrence of certain diseases, specifically coronary diseases. The correlation appears to be weaker when negative affect is considered (that is, the search for a correlation between a higher negative affect and higher incidence of diseases is inconclusive). Furthermore, the authors observe that they are "not able to draw strong conclusion about which specific types of feelings are most related to health". Indeed, the interesting aspect in this area is that the number of (longitudinal) studies is very high, but the results are sometimes conflicting. Cohen and Pressman (Cohen & Pressman, 2006) provide a review of studies with a particular focus on elderly population. A very interesting finding in this case is that a correlation has been found between health and happiness in independently living older adults, but not for adults living in assisted living structures.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

One could argue that Exergames provide a great opportunity for people living independently at home to come together, either by attending a game session in person or by attending online from home, and exercise, and socialize, and hopefully have a good time and become happier. Besides being beneficial in its own right, happiness (and, specifically, frequent positive affect) could reduce the occurrence of certain deceases. Exergames is a facilitator, a methodology to achieve these positive effects and to help motivate elderly to perform these activities.

SOCIALISATION, MOTIVATION, AND PERSUASIVE TECHNIQUES

There are a lot of usability studies for exergames for senior users (Gerling et al., 2010; Ijsselsteijn, Nap, de Kort, & Poels, 2007; Pinelle, Wong, & Stach, 2008; Wüest et al., 2014), however the usability itself cannot be considered a sufficient motivation to use the game (Ijsselsteijn et al., 2007).

There is evidence that both physical activity, social contacts and cognitively stimulating activity may improve cognitive performance among elderly (Gallucci et al., 2009). There is also general consensus that all physical activity have considerable effect on the physical decline related to ageing, and that active ageing will have very positive effects both for the society and for the individual.

While it is hard to argue that socialisation directly leads to better physical health, socialisation is a strong motivating factor in following exercise programs. This might work by people being motivated by simply enjoying being together with other people, but they may also be motivated by others seeing them participate, or by not wanting them to see that they are not participating. The Nettrim and GameUp projects (S. T. Konstantinidis et al., 2014) try to closer examine all these sides of socialisation as a motivating factor, and by understanding them more closely, try to take advantage of this in designing motivating exergames.

Ijsselsteijn (Ijsselsteijn et al., 2007) identified four main areas for design opportunities for gaming targeting elderly including socializing with others within and outside their social network. Social interaction in games is being considered the main motivation for elderly users that also foster the intergeneration games (Khoo, Cheok, Nguyen, & Pan, 2008). A number of game projects showed that socialisation and communication (Lin, Mamykina, Lindtner, Delajoux, & Strub, 2006; Shim, Baecker, Birnholtz, & Moffatt, 2010; Toscos, Faber, Connelly, & Upoma, 2008) while playing or performing physical activity (Festinger, 1954) is more likely to achieve the desired goal. On the other hand, (Ellen Brox et al., 2011) stated that the social factor is really important regarding exergames for seniors, but seniors are sceptic, since they worried of the replacement of physical meetings with online ones. To this extent, (Gajadhar et al., 2010) show that online co-playing was deemed less rewarding than physical co-playing or when the seniors thought they were playing against a computer.

Persuasion can be summed up in the following buzzwords: behavioural change, attitude change, motivation, change in worldview and compliance (B J Fogg, 2003), and similarly, persuasion defined as the deliberate use of communication to change attitudes and behaviour of people. (Chatterjee & Price, 2009). To this extent, "Captology" an acronym based on the phrase: "Computers as Persuasive Technologies" is a term that coined by (B J Fogg, 2003), *"focuses on the design, research and analysis of interactive computing products created for the purpose of changing people's attitudes or behaviours"*. Persuasive technologies are part of people's everyday life (P. King & Tester, 1999) and they should not be left out from exergames for seniors, since in specific context persuasive technologies can play a positive role by convincing stimulating or motivating users to engage in health behaviours (Kort et al., 2005).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Designing and implementing persuasive technologies is not an easy task. According to (Bj J Fogg, 2009) many failures might occur before creating a successful product. Oinas-Kukkonen and Harjumaa proposed the Persuasive System Design (PSD) model, a framework for designing and evaluating persuasive systems, in which twenty-eight design guidelines were defined (Oinas-Kukkonen & Harjumaa, 2009). An eight-step design process in the early stages of persuasive technologies (mainly focusing on behaviour change) has been provided by (Bj J Fogg, 2009) to serve as best practices:

1. Choose a simple behaviour to target;
2. Choose a receptive audience (where possible);
3. Find what prevents the target behavior;
4. Choose a familiar technology channel;
5. Find relative examples of persuasive technology;
6. Imitate successful examples;
7. Test and iterate quickly;
8. Expand on success.

In addition persuasive strategies should be part of the design process to create persuasive technologies. Steps 3 to 6 could be analysed, identified and implemented easier through persuasive strategies, increasing the success of the final product. Persuasive strategies vary according to the users, the technology channel and the behaviour to target. A set of persuasive strategies for seniors increasing their physical activity through games identified by (Ellen Brox et al., 2011): display information to encourage people to be more active; record and display the user's past behaviour; use positive reinforcements to improve behaviours; Make an attractive and friendly user interface; provide information at opportune moments; use social influence.

Those strategies could be fulfilled by "personalizing", a strategy proposed by (P. King & Tester, 1999) as a general strategy. While all physical activities have positive health effects, there are few physical activities that are suited for all people. Activities need to be adjusted to the individual. At the same time, the motivating factors, needs to be individually adjusted. This individualized communication, is usually referred to as 'tailoring'. (Noar, Benac, & Harris, 2007)

Hawkins (Hawkins, Kreuter, Resnicow, Fishbein, & Dijkstra, 2008) divides tailoring strategies into personalization, feedback and content matching. Though the strategies are usually combined, they do constitute unique conceptual ideas. Personalization attempts to increase attention or motivation to process the conveyed messages through for instance techniques like 'identification', 'raising expectations' or 'contextualization'. Feedback presents to the individuals their own or their competitors/collaborators health status or their activities. Content matching tries to give the user relevant information on the behaviour of interest.

A MOTIVATIONAL FRAMEWORK: THE CASE OF THE GAMEUP AND NETRRIM PROJECTS

A framework was developed based on the existing aforementioned theories for motivation and persuasive technologies. Based on the input of the users performance a selection of predefined messages

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

were made and provided to the user screen at the appropriate time or game events. The Motivational Engine includes decision making techniques to select the appropriate motivational message regarding the timeframe of the game.

The exergame is an apple picking game (Figure 1) using a motion camera to capture the movements of the senior while catching the falling apples and placing them in the correct basket. The goal of the users is to catch as many apples as possible in a predefined time (different for each user). The game aims to improve the balance of the seniors.

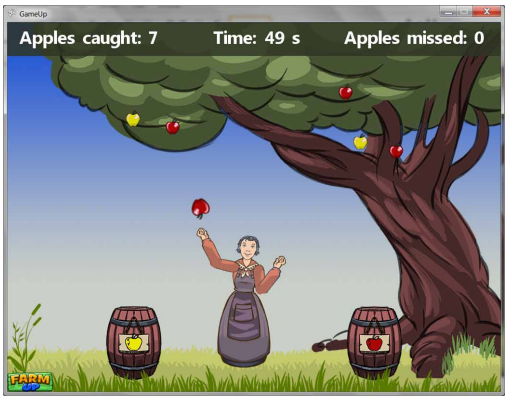
The framework gets both personal and team goals as input divided in 3 subcategories each.

For personal goals, the first category is the goal that the user is supposed to reach, the second is the current state of the user and the third is the results that are achieved by the user.

In addition the framework input also includes experts' predefined messages as input: "messages for personal progress" (during the game) and "messages for personal achievements" (at the end of the game). The messages for personal progress are positive and divided into 4 categories regarding the user performance: excellent progress; above average progress; average progress; and below average progress. As average progress either the average efforts of user record in the database or the average efforts of users with similar criteria (age, illnesses, abilities, etc.) could be defined. In the same way the positive messages for personal achievements are divided into 4 similar categories: excellent, above average, average and below average achievements. An average progress can be defined either as the average efforts of user recorded in the database to achieve the same goal or the average efforts of users recorded in the database to achieve the same goal with similar criteria (age, illnesses, abilities, etc.).

Two categories of output are provided by the framework. Firstly the "Personal Progress" category which is the feedback to the user during the game and this includes: a bar or a number that states the current progress of the game regarding the goal; a bar or a number that states the time that has passed and the remaining time; and predefined messages from experts. Secondly, the "Personal Achievements" category which is depicted after the end of the game and a combination of the following types of information could be depicted to the user: reach the goal message from experts; an image/video/animation represents how close to the goal the user is; the final score and the time spent.

Figure 1. The exergame: apple picking game



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

For Team goals the first category of input is the collaborative goal that the team is intent to reach, the second is the current state of the team goal and the third is the results of users that is achieved by the team.

The positive predefined messages from the experts as input to the framework are also divided into two categories: progress ("Messages for team progress") and achievements ("Messages for Team Achievements"), but there are also "Messages for personal progress within the team progress" and "Messages for personal achievements within the team achievements". All the messages are divided into 4 categories regarding the user and the team progress and achievements. Comparison of the team progress or achievements (average progress or average achievements) could be defined either as the average effort of teams recorded in the database or the average efforts of teams recorded in the database with similar criteria (age, illnesses, abilities, etc.) and similarly for the achievements. When it comes to "personal progress within the team goals", the messages are based on the average efforts of players compared with other members of the team or an average effort by players with similar criteria (age, illness, abilities, etc.) in the same team. The messages for "personal achievements within the team achievements" are defined respectively.

The output in this case includes the personal achievements, but also the team results. Two categories of output are provided by the framework for team goals, but personal goals could also have been used. In addition user contribution to the team is emphasized and positive messages are used. The "Team Progress" gives the feedback to the user about the progress of the team using: a bar or a number states the current progress of the team game regarding the goal; a bar or a number states the time that has passed and the remaining time; and predefined messages from experts are given. The "Team Achievements" category which is shown after the end of the game, depicts a combination of the following types of information: reached the team goal message from experts; an image/video/animation represents how close to the team goal the team is; the team's final score and the time; the user's final score and the time; and a comparison indicator for the user regarding the team.

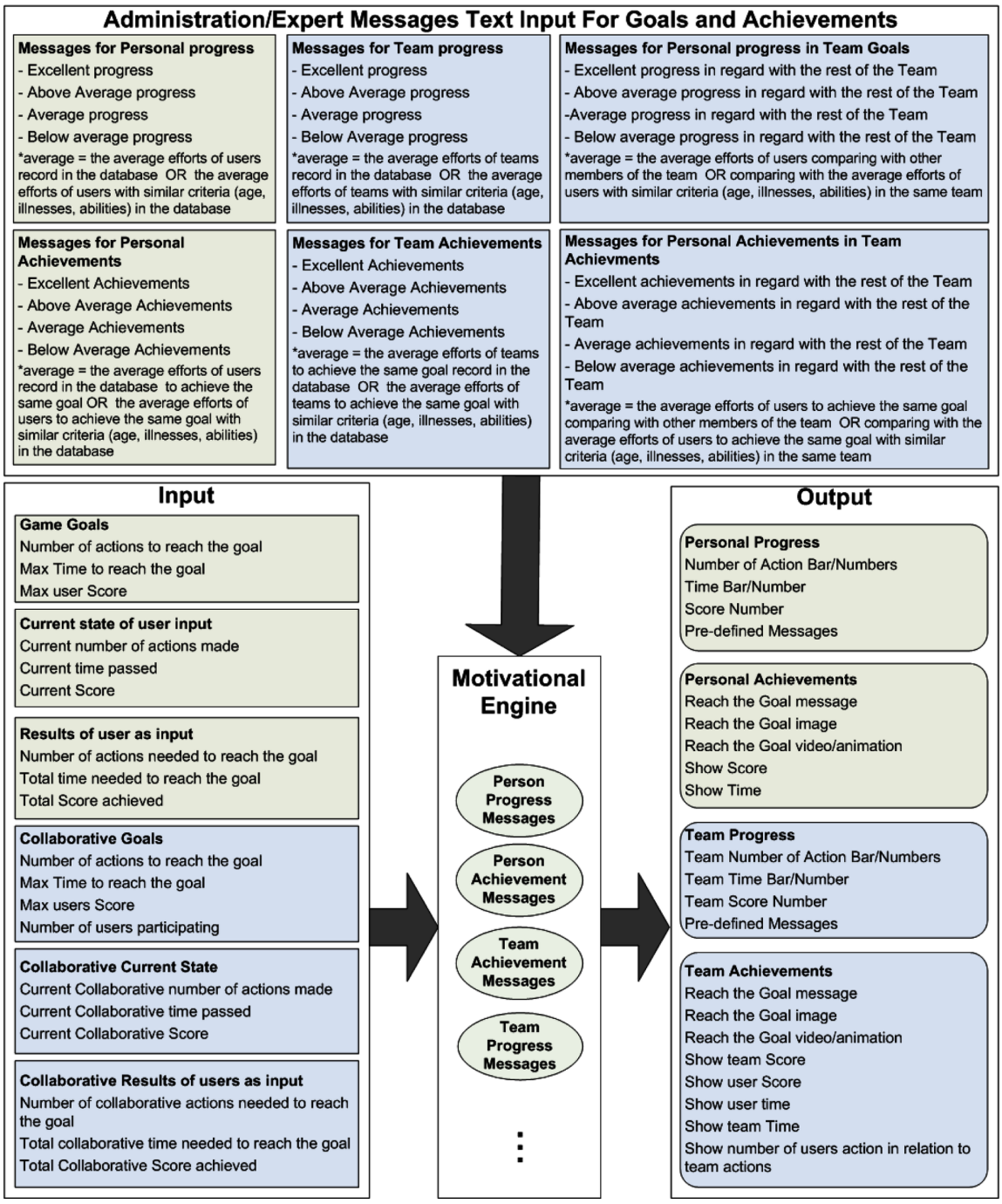
The "Motivational Engine" is responsible for suitable output by utilising appropriate algorithms on the input. Figure 2 shows an abstract of the architecture of the Motivational framework.

The motivational framework presented is following the persuasive strategies below and need an online game to be applied.

- *Display information to encourage people to be more active.* Both personal and team progress could be provided to the user in order for her to know her current state of progress. In addition, motivational messages encourage her to be more active.
- *Record and display the user's past behaviour.* The user's and the team's past behaviour is being recorded, displayed and compared with the latest one. The messages from experts are selected according to the user's past behaviour and performance.
- *Use positive reinforcements to improve behaviours.* The messages from the experts are all formed in a positive style in order to improve behaviour and performance. Even the messages for below average progress or achievement focus on next efforts of the user/team and guide them to increase the participation and performance.
- *Make an attractive and friendly user interface.* Despite the fact that the framework itself does not provide an interface, the elements (bars or numbers and messages) which it proposes can be used within an attractive and user-friendly interface without providing hazard to it.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Abstract architecture of the motivational framework



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- *Provide information at opportune moments.* The framework itself cannot determine when the output for progress will be presented, but gives the opportunity to be presented continuously, or at a chosen time. The achievements output should be presented after the end of the session or the game.
- *Use social influence.* The team goals and the team messages which can be set within this framework increase the social influence. Seniors interact with other seniors and try together to achieve a common goal. The element of comparison between seniors is minimized and is used as positive reinforcement to encourage seniors to be more active.
- *Personalization.* The goals set for each user are personalized according to her abilities. In addition, the goals set for each team are a sum of user goals according to each user's abilities. The comparison that takes place with other users/teams in order to provide a positive message is taking place either with the user's previous performance or with a user with similar criteria (age, illnesses, abilities, etc.).

Within the limitation of the aforementioned framework, the input of the experts is the most crucial. Positive messages should be tested and adjust to the seniors need in each different environment and game.

Socialisation has been proven to be an important factor of everyday life, especially for seniors. In this section existing efforts on the inclusion of this aspect into exergames have been described. Motivation on the other hand is really important so as to keep the seniors play the exergame frequently. Persuasive techniques that are used in exergaming are described and the connection with the social aspects emphasized. A motivational framework was presented as a case which implements the persuasive strategies for seniors found in literature.

FUTURE RESEARCH

To be able to provide exercise applications to a large number of elderly in a cost effective way, the logical choice is to provide them through a mobile platform, either on smart phones or on tablets. This can either be applications that use sensors in the phones/tablets, like movement or positioning, or it can be simple applications that only encourage the users to exercise and have means of communicating between users. This approach can however be challenging due to the screen size, and particularly phone screens will be too small for most gaming and gamification purposes. Also if Kinect is being used, a big screen is needed since the players have to be at a certain distance from the camera.

One aim for future trials could be to get new knowledge about what effect exergames could have for elderly, or to get knowledge about how exergames can be designed to improve any positive effects. Clinical protocols and standards should be formed and validated in order to measure the effect of exergames on seniors with clinical trials and provide the evidence that the clinical practice requires.

Most studies of exergames have a clinical effect component, even if in some cases they are just measuring whether the system is used – assuming all activity is positive. This is natural since the reason for using the games is to improve health and wellbeing. To use a standardised tool for testing effect, gives advantages in both reliability and validity. Within the field of exercises for elderly, there are several well tested tools for measuring effect.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In general these tools are considered to give more valid and reliable results, but the downside is that the tools usually require a physician to be together with the patient while performing the test. The alternative would be to use sensors and to rely on direct measures like step-counting, O2-consumption and heart rate. Changes in resting heart rate can for instance be used as a measurement of fitness.

It is also possible to design studies where the goal is not to get a better health effect than ordinary training. The study can be monitored for not having negative effects, while the real benefit might be in improved availability. Such a study can be performed by doing only an economic analysis of the alternate value.

Another approach is to assume that physical activity by definition will give health benefits, and instead of evaluating the health effect, you can monitor the use and/or compliance. In most cases this is significantly easier to measure than to measure the user's health status, though it might be harder to publish such studies in medical journals, since the clinical outcome is only measured through proxy. There are however journals within several other fields that are accepting such studies.

Instead of focusing on what effects it has on the user to use specific software, it is also possible to focus on how a specific piece of software can be improved to give better effects. At least two aspects are relevant for a study: How can the usability be improved, and how can the software be more motivating for the user.

Usability of software is a popular topic for many studies with a technological approach. There are multiple standardised tools for measuring usability, some of the most popular are:

- Software Usability Measurement Inventory (SUMI);
- Website Analysis and Measurement Inventory (WAMI);
- System Usability Scale (SUS);
- Questionnaire for User Interaction Satisfaction (QUIS);
- Usefulness, Satisfaction, and Ease of use Questionnaire (USE);
- Computer System Usability Questionnaire (CSUQ);
- Isometric Usability Inventory (IsoMetric).

What a future project should concentrate on will among other things depend on the participants. For instance testing of the physical effects of exergames requires that health care organisations are actively taking part, while testing of design and usability are more technical issues.

CONCLUSION

In this chapter the current trends of games for physical training that are used by seniors are described. Notions of exergames, gamified applications and online social games are analysed in order to provide a current state-of-the-art for a rapidly expanding research area, but also for new markets. In general exercise games for seniors are well accepted by them (Chao et al., 2013; Williams et al., 2010). However, there are a lot of design parameters that should be improved to match the usability needs of the seniors. Scores and progress measurements are too generic for rehabilitation use (Anderson et al., 2010), many of the games are difficult to play (Anderson et al., 2010) and many games are too fast (Theng et al., 2009; Wollersheim & Merkes, 2010).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The effect of games for physical activity for seniors has not yet been proven and existing research needs further expansion, but there are no indications that exergames are worse than other training programs. (Bleakley et al., 2013; Larsen et al., 2013; Webster & Celik, 2014) have registered a number of attempts to compile evidence of the effect of exergames for seniors. However these studies are not focusing on online social exergames for seniors. There is thus an urgent need of a definition of clinical evaluation frameworks in order to get the required evidence for the effects on the quality of life of seniors that arise from online exergames.

A big number of current and past projects reveal the intention of the society of further research for physical training of seniors through games. The existing projects use biomedical sensors (acelometers, custom movement systems, heart rates, etc.) and wearble sensors in order to capture the users' movements, while others just use a phone app to inform and motivate the users. The latest project advancements engage the use of games that have a game engine and are controlled with a motion capture camera or motion controls. A few of the projects combine both gamified applications and exergames. To this extent, online aspects of interaction between users and within the game are not common, but social interactions have been considered an important aspect, though so far they are mainly implemented through "traditional" communication technologies. The exergames for seniors need to be personalised and a lot of existing efforts through research project are focusing on that.

The technology is rapidly changing and the reseach for games for physical training of seniors is changing as well. New motion tracking devices are continusly appearing on the market and allow the creation of different results, covering different user needs. A sustainable game should be adaptive, and use the motion tracking device that really is needed, without overloading a senior user with uncomfortable sensors.

Exergames for seniors provide the necessary platform for people to meet and share a good and social time together, and to help them motivate themselves and each other to stay physically active. Despite the fact that exercising does not only have a direct and even immediate positive impact on brain health, it also gives us a release from stress and gives us a way to expend that enormous surplus of calories that many of us ingest on a daily basis. Furthermore exergames provide a great opportunity for people living independently at home to come together, either by attending a game session in person or by attending online from home, and have exercise, and socialize, and hopefully have a good time and become happier (Kaplan et al., 2013). The special needs of seniors should always be taken into account since online soacial games sometimes could be seen as a threat for inperson social interaction (Ellen Brox et al., 2011).

It is hard to argue that socialisation directly leads to better physical health, but socialisation is a strong motivating factor in exercise programs. In addition persuasive technologies can play a positive role by convincing stimulating or motivating users to engage in health behaviours. A persuasive strategy for seniors within games for physical training has started to be developed. The guidelines identified in the current chapter and a best practice of a generalized motivation framework is presented herein. Facilitators' input in that framework could limit the effectiveness if positive feedback is not taken into consideration. However, there is still a need for further research on motivational and persuasive strategies regarding social exergames for seniors and best practice guidelines in the form of a common standard are needed.

Exergames and mHealth applications for psysical training are on the starting line. Cost-effective solutions will likely emerge over the next years and further studies regarding clinical effects are needed. Usability should always be taken into consideration, especially the seniors' needs and abilities.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Having all that in mind, one could consider use of information and communication technologies (ICT) by seniors. Despite the fact that seniors increase their digital literacy in order to perform everyday transactions (e.g. bill payments, bank transactions, etc.), there is still a question regarding the similarity of those actions with the use of exergames and gamified applications for physical training.

The increase of "older" citizens in many countries will limit the resources for care in the future. The solution is believed to involve an active and healthy ageing. Seniors could prolong their independent living and health by personalized ICT services. The existing research reveals that the emerging area of games for keeping seniors active and healthy in the form of exergames and gamified application in combination with sensors and motion devices contributes to the solution of keeping the seniors healthy and active. Formal carers have started to experiment with the existing exergames with initial positive impact. Thus, online social exergames for seniors could acts as a pillar of gamification for clinical practice.

ACKNOWLEDGMENT

This work is partially supported by "Nettrim: Social online exercises for elderly" (http://nettrim.norut.no), a project co-funded under the Interreg IV A programme and co-funded by Troms Fylkeskommune (Troms fylkeskommune: RUP2013-0014 / Interreg IVA Nord 2013-0021) and by "GameUp: Game-based mobility training and motivation of senior citizens", a project co-funded by the AAL Joint Programme (REF. AAL-2011-4-090), the Research council of Norway, the Federal Department of Economic Affairs, Education and Research EAER - State Secretariat for Education, Research and Innovation SERI of Switzerland and the Ministering of Industry, Energy and Tourism of Spain. The authors would like to thank the participants in the studies and the partners of the projects.

REFERENCES

Alaoui, M., Ting, K. L. H., & Lewkowicz, M. (2014). The Urge for Empirically-informed Design of Social-Oriented AAL Applications – The Example of 2 AAL Projects. *Procedia Computer Science*, *27*, 379–388. doi:10.1016/j.procs.2014.02.042

Anderson, F., Annett, M., & Bischof, W. F. (2010). Lean on Wii: Physical rehabilitation with virtua reality Wii peripherals. *Studies in Health Technology and Informatics*, *154*, 229–234. PMID:20543303

ASEAN Strategic Framework for Social Welfare and Development. (2012The ASEAN Secretariat. Jakarta.

Bamidis, P., Frantzidis, C., Kyrillidou, A., Ladas, A., Grigoriadou, E., Billis, A., & ... (2011). Cognitive training, physical exercise and information technology: Neuroscientific challenges and first evidence from the LLM project. *Neuroscience Letters*, *500*, e6. doi:10.1016/j.neulet.2011.05.080

Bamidis, P. D., Konstantinidis, E. I., Billis, A., Frantzidis, C., Tsolaki, M., & Hlauschek, W. … Pattichis, C. S. (2011). A web services-based exergaming platform for senior citizens: the Long Lasting Memories project approach to e-health care. Proceedings of the *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 2505–2509).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Baranowski, T., Buday, R., Thompson, D. I., & Baranowski, J. (2008). Playing for Real. Video Games and Stories for Health-Related Behavior Change. *American Journal of Preventive Medicine*. PMID:18083454

Berry, L. L., & Bendapudi, N. (2007). Health Care: A Fertile Field for Service Research. *Journal of Service Research*, *10*(2), 111–122. doi:10.1177/1094670507306682

Billis, A. S., Konstantinidis, E. I., Mouzakidis, C., Tsolaki, M. N., Pappas, C., & Bamidis, P. D. (2010). A game-like interface for training seniors' dynamic balance and coordination. *IFMBE Proceedings*, *29*, 691–694. doi:10.1007/978-3-642-13039-7_174

Bleakley, C. M., Charles, D., Porter-Armstrong, A., McNeill, M. D. J., McDonough, S. M., & McCormack, B. (2013). Gaming for Health: A Systematic Review of the Physical and Cognitive Effects of Interactive Computer Games in Older Adults. *Journal of Applied Gerontology*, 0733464812470747–0733464812470747–. doi:10.1177/0733464812470747

Healthy aging through healthy living: towards a comprehensive policy and planning framework for seniors in B.C.: a discussion paper. (2005). British Columbia. Ministry of Health. Retrieved from http://www.health.gov.bc.ca/library/publications/year/2005/healthy_aging.pdf

Brox, E., Evertsen, G., Horst-Fraile, S., Browne, J., & Åsheim-Olsen, H. (2015). Experience with a 3D Kinect exergame for elderly. Proceedings of the *8th International Conference on Health Informatics -Healthinf2015 [submitted]*, Lisbon, Portugal.

Brox, E., Luque, L. F., Evertsen, G., & Hernandez, J. E. G. (2011). Exergames for elderly: Social exergames to persuade seniors to increase physical activity. In *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops* (pp. 546–549). Dublin: IEEE.

Cameirão, M. S., Badia, S. B. I., Oller, E. D., & Verschure, P. F. M. J. (2010). Neurorehabilitation using the virtual reality based Rehabilitation Gaming System: Methodology, design, psychometrics, usability and validation. *Journal of Neuroengineering and Rehabilitation*, *7*(1), 48. doi:10.1186/1743-0003-7-48 PMID:20860808

Cantwell, D., Broin, D. O., Palmer, R., & Doyle, G. (2012). Motivating Elderly People to Exercise Using a Social Collaborative Exergame with Adaptive Difficulty. *Proceedings of the European Conference on Games Based Learning* (p. 615).

Carús, J. L., García, S., García, R., Waterworth, J., & Erdt, S. (2014). The ELF@Home project: Elderly sELF-care based on sELF-check of health conditions and sELF-fitness at home. *Studies in Health Technology and Informatics*, *200*, 164–166. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24851984 PMID:24851984

Centers for Disease Control and Prevention. (2013). *The State of Aging and Health in America 2013*. Atlanta, GA. Retrieved from http://www.cdc.gov/features/agingandhealth/state_of_aging_and_health_in_america_2013.pdf

Chao, Y. Y., Scherer, Y. K., Wu, Y. W., Lucke, K. T., & Montgomery, C. A. (2013). The feasibility of an intervention combining self-efficacy theory and Wii Fit exergames in assisted living residents: A pilot study. *Geriatric Nursing*, *34*(5), 377–382. doi:10.1016/j.gerinurse.2013.05.006 PMID:23764366

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Charter of fundamental rights of the European Union (2010/C 83/02). (2010). *Official Journal of the European Union, C 83/389*.

Chatterjee, S., & Price, A. (2009). Healthy Living with Persuasive Technologies: Framework, Issues, and Challenges. *Journal of the American Medical Informatics Association*, *16*(2), 171–178. doi:10.1197/jamia.M2859 PMID:19074300

Cohen, S., & Pressman, S. D. (2006). Positive Affect and Health. *Current Directions in Psychological Science*, *15*(3), 122–125. doi:10.1111/j.0963-7214.2006.00420.x

De Kort, Y. A. W., IJsselsteijn, W. A., Eggen, J. H., & Van den Hoven, E. A. W. H. (2005). Persuasive gerontechnology. *Gerontechnology (Valkenswaard)*, *4*(3), 123–127. doi:10.4017/gt.2005.04.03.001.00

Decision No 1312/2013. (2013, December 11). The European parliament and of the council of 2013 on the Strategic Innovation Agenda of the European Institute of Innovation and Technology (EIT): the contribution of the EIT to a more innovative Europe. *Official Journal of the European Union, L 347/892*.

Diener, E., & Chan, M. Y. (2011). Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity. *Applied Psychology: Health and Well-Being*, *3*(1), 1–43.

DOSSy:Digital Outdoor and Safety System [project website]. (2012). Retrieved from http://dossy-aal.iwi.unisg.ch/index.php?id=the-project

Dubowsky, S., Genot, F., Godding, S., Kozono, H., & Skwersky, A. (2000). PAMM - a robotic aid to the elderly for mobility assistance and monitoring: a "helping-hand" for the elderly. *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)* (Vol. 1, pp. 570–576). IEEE. 10.1109/ROBOT.2000.844114

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. doi:10.1177/001872675400700202

Fogg, B. J. (2003). Persuasive technology: using computers to change what we think and do. In Y. Kort, W. IJsselsteijn, C. Midden, B. Eggen, & B. J. Fogg (Eds.), Persuasive Technology Using Computers to Change What We Think and Do (Vol. 5, p. 283). Morgan Kaufmann. doi:10.4017/gt.2006.05.01.009.00

Fogg, B. J. (2009). Creating Persuasive Technologies : An Eight-Step Design Process. *Technology (Elmsford, N.Y.)*, *91*, 1–6. doi:10.1145/1541948.1542005

Gajadhar, B. J., Nap, H. H., De Kort, Y. A. W., Ijsselsteijn, W. A., Gajadhar, B. J., & Nap, H. H., … Ijsselsteijn, W. A. (2010). Out of Sight, out of Mind : Co-Player Effects on Seniors' Player Experience. *Proceedings of the Fun and Games Conference* (pp. 74–83). 10.1145/1823818.1823826

Gallucci, M., Antuono, P., Ongaro, F., Forloni, P. L., Albani, D., Amici, G. P., & Regini, C. (2009). Physical activity, socialization and reading in the elderly over the age of seventy: What is the relation with cognitive decline? Evidence from "The Treviso Longeva (TRELONG) study." *Archives of Gerontology and Geriatrics*, *48*(3), 284–286. doi:10.1016/j.archger.2008.02.006 PMID:18374429

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Gerling, K. M., Schild, J., & Masuch, M. (2010). Exergame design for elderly users: the case study of SilverBalance. *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology - ACE '10* (pp. 66–69). ACM Press. 10.1145/1971630.1971650

Goldspink, D. F. (2005). Ageing and activity: Their effects on the functional reserve capacities of the heart and vascular smooth and skeletal muscles. *Ergonomics*, *48*(11-14), 1334–1351. doi:10.1080/00140130500101247 PMID:16338704

Gotsis, M., Wang, H., Spruijt-Metz, D., Jordan-Marsh, M., & Valente, F. T. W. (2013). Wellness Partners: Design and Evaluation of a Web-Based Physical Activity Diary with Social Gaming Features for Adults. *JMIR Res Protoc*, *2*(1), e10. doi:10.2196/resprot.2132 PMID:23611986

Goudie, R. J. B., Mukherjee, S. N., Sr., De Neve, J.-E., Oswald, A. J., & Wu, S. (2011). Happiness as a Driver of Risk-Avoiding Behavior. *CESifo Working Paper Series No. 3451*. Retrieved from http://ssrn.com/abstract=1846390

Hallberg, J., Casati, F. (2013). *Technical Foresight Report on Mental Wellbeing for Active Healthy Ageing*. To be published by EIT ICT Labs Health and Wellbeing.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. *Proceedings of System Sciences (HICSS), 2014 47th Hawaii International Conference* (pp. 3025–3034). doi:10.1109/HICSS.2014.377

Hawkins, R. P., Kreuter, M., Resnicow, K., Fishbein, M., & Dijkstra, A. (2008). Understanding tailoring in communicating about health. *Health Education Research*, *23*(3), 454–466. doi:10.1093/her/cyn004 PMID:18349033

Heller, B., Senior, T., & Wheat, J. (2014). The Smartfloor: A Large Area Force-measuring Floor for Investigating Dynamic Balance and Motivating Exercise. *Procedia Engineering*, *72*, 226–231. doi:10.1016/j.proeng.2014.06.040

Honka, A., Kaipainen, K., Hietala, H., & Saranummi, N. (2011). Rethinking health: ICT-enabled services to empower people to manage their health. *IEEE Reviews in Biomedical Engineering*, *4*, 119–139. doi:10.1109/RBME.2011.2174217 PMID:22273795

Hors Fraile, S., Browne, J., Brox, E., & Evertsen, G. (2012). Suitability analysis of commercial open-source driven motion sensor devices applied to exergames for the elderly. In AAL forum. Eindhoven.

Ijsselsteijn, W., Nap, H. H., de Kort, Y., & Poels, K. (2007). Digital Game Design for Elderly Users. *Proceedings of the 2007 Conference on Future Play* (pp. 17–22). New York, NY, USA: ACM. 10.1145/1328202.1328206

Johnsen, E., Burkow, T., & Vognild, L. K. (2012). Den sosiale gruppa som motivasjon til fysisk trening. In Helsesosiologi: Analyser av helse, sykdom og behandling (pp. 325–346). Gyldendal Akademisk.

Jung, Y., Li, W., Gladys, C., & Lee, K. M. (2009). Games for a Better Life : Effects of Playing Wii Games on the Well-Being of Seniors in a Long-Term Care Facility. *Proceedings of the Sixth Australasian Conference on Interactive Entertainment* (pp. 0–5). 10.1145/1746050.1746055

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kaplan, S., Madden, V. P., Mijanovich, T., & Purcaro, E. (2013). The perception of stress and its impact on health in poor communities. *Journal of Community Health*, *38*(1), 142–149. doi:10.100710900-012-9593-5 PMID:22806256

Karlhuber, I., & Schürz, S. (2013). MOBILE.OLD [poster]. In *Geriatric Congress*. Vienna (Austria). Retrieved from http://www.bluepoint-it.ro/mobiledotold/wp-content/uploads/2013/07/MdO_URA_Poster_v1.2.pdf

Khoo, E. T., Cheok, A. D., Nguyen, T. H. D., & Pan, Z. (2008). Age invaders: Social and physical inter-generational mixed reality family entertainment. *Virtual Reality (Waltham Cross)*, *12*(1), 3–16. doi:10.100710055-008-0083-0

King, D., Greaves, F., Exeter, C., & Darzi, A. (2013). "Gamification": Influencing health behaviours with games. *Journal of the Royal Society of Medicine*, *106*(3), 76–78. http://www.ncbi.nlm.nih.gov/pubmed/23481424 doi:10.1177/0141076813480996 PMID:23481424

King, P., & Tester, J. (1999). The Landscape of Persuasive Technologies. *Communications of the ACM*, *42*(5), 31–38. doi:10.1145/301353.301398

Konstantinidis, E. I., Billis, A., Hlauschek, W., Panek, P., & Bamidis, P. D. (2010). Integration of cognitive and physical training in a smart home environment for the elderly people. *Studies in Health Technology and Informatics*, *160*, 58–62. PMID:20841650

Konstantinidis, S. T., Fernández-Luque, L., Brox, E., & Kummervold, P. E. (2014). Seniors Exergaming 2.0: The role of social and motivational aspects on games for elderly's physical training through web 2.0 techniques. Proceedings of Medicine 2.0 conference, Malaga, Spain.

Larsen, L. H., Schou, L., Lund, H. H., & Langberg, H. (2013). The Physical Effect of Exergames in Healthy Elderly—A Systematic Review. *Games for Health Journal*, *2*(4), 205–212. doi:10.1089/g4h.2013.0036 PMID:26192224

Lauritzen, J., Muñoz, A., Luis Sevillano, J., & Civit, A. (2013). The usefulness of activity trackers in elderly with reduced mobility: A case study. *Studies in Health Technology and Informatics*, *192*(1-2), 759–762. PMID:23920659

Lin, J., Mamykina, L., Lindtner, S., Delajoux, G., & Strub, H. (2006). Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. In P. Dourish & A. Friday (Eds.), UbiComp 2006: Ubiquitous Computing SE - 16 (Vol. 4206, pp. 261–278). Springer Berlin Heidelberg. doi:10.1007/11853565_16

Maggiorini, D., Ripamonti, L. A., & Zanon, E. (2012). Supporting seniors rehabilitation through videogame technology: A distributed approach. In *2012 2nd International Workshop on Games and Software Engineering: Realizing User Engagement with Game Engineering Techniques, GAS 2012 - Proceedings* (pp. 16–22). 10.1109/GAS.2012.6225920

Massé, F., Paraschiv-Ionescu, A., Ženko, B., Džeroski, S., & Aminian, K. (2013). Lifestyle evaluation using wearable technologies: opportunities for stroke patients. In Converging Clinical and Engineering Research on Neurorehabilitation (pp. 941–945). Springer. doi:10.1007/978-3-642-34546-3_153

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

McCallum, S. (2012). Gamification and serious games for personalized health. *Studies in Health Technology and Informatics*, *177*, 85–96. PMID:22942036

Meurer, J., & Wieching, R. (2012). Motivating Elderly People to Use Fall Preventive Exercise Training Games at Home: Are Community Based ICT Features Always a Good Choice? In S. Budweg, C. Müller, M. Lewkowicz, V. Pipek, & M. Rohde (Eds.), *Designing for Inter/Generational Communities - 3rd International Workshop "Fostering Social Interactions in the Ageing Society", Proceedings of COOP Conference 2012* (p. 28). IISI - International Institute for Socio-Informatics.

Miller, A. S., & Cafazzo, J. a, & Seto, E. (2014). A game plan: Gamification design principles in mHealth applications for chronic disease management. *Health Informatics Journal*, *1*(Jul). doi:10.1177/1460458214537511 PMID:24986104

Miyosawa, T., Yanagisawa, S., & Okuhara, M. (2013). Development and Evaluation of a Rehabilitation Program using Kinect ™ Motion Capture Technology. Proceedings of *CENTRIC 2013 : The Sixth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services* (pp. 37–43). Venice, Italy. Retrieved from http://www.thinkmind.org/download.php?articleid=centric_2013_3_10_30029

momo: Mobility Motivator [project webpage]. (2012). Retrieved from http://www.mobility-motivator.uvsq.fr/

Nawaz, A., Waerstad, M., Omholt, K., Helbostad, J. L., Vereijken, B., Skjaeret, N., & Kristiansen, L. (2014). Designing Simplified Exergame for Muscle and Balance Training in Seniors : A Concept of "Out in Nature." Proceedings of the *8th International Conference on Pervasive Computing Technologies for Healthcare*. Oldenburg, Germany: ACM Digital Library. 10.4108/icst.pervasivehealth.2014.255269

Nguyen, T., Tapanainen, T., & Obi, T. (2014). A review of information and communication technology (ict) training for elderly people – toward recommendations for developing countries. *Proceedings* of *PACIS 2014* (Paper 267). Retrieved from http://aisel.aisnet.org/pacis2014/267

Noar, S. M., Benac, C. N., & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673–693. doi:10.1037/0033-2909.133.4.673 PMID:17592961

Novak, D., Nagle, A., & Riener, R. (2014). Can two-player games increase motivation in rehabilitation robotics? *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI '14* (pp. 447–454), New York, USA. ACM Press. 10.1145/2559636.2559658

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive Systems Design: Key Issues, Process Model, and System Features. *Communications of the Association for Information Systems*, *24*(1), 485–500. Retrieved from http://www.few.vu.nl/~wissen/downloads/seminar/2011_Oinas-kukkonen.pdf

Olphert, W., & Damodaran, L. (2013). Older people and digital disengagement: A fourth digital divide? *Gerontology*, *59*(6), 564–570. doi:10.1159/000353630 PMID:23969758

Otten, R. (2012). Go Dibiking [presentation]. Proceedings of the *AAL Forum 2012*. Eindhoven, The Netherlands.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Pinelle, D., Wong, N., & Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. *Proceedings of the SIGCHI Conference on human factors in computing systems*, 1453–1462. 10.1145/1357054.1357282

Reiss, A., & Stricker, D. (2011). Introducing a modular activity monitoring system. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (pp. 5621–5624).

Reiss, A., Stricker, D., & Lamprinos, I. (2012). An Integrated mobile system for long-term aerobic activity monitoring and support in daily life. *Proc. of the 11th IEEE Int. Conference on Trust, Security and Privacy in Computing and Communications, TrustCom-2012 - 11th IEEE Int. Conference on Ubiquitous Computing and Communications, IUCC-2012* (pp. 2021–2028). 10.1109/TrustCom.2012.81

Rendon, A. A., Lohman, E. B., Thorpe, D., Johnson, E. G., Medina, E., & Bradley, B. (2012). The effect of virtual reality gaming on dynamic balance in older adults. *Age and Ageing*, *41*(4), 549–552. doi:10.1093/ageing/afs053 PMID:22672915

Renehan, E., Dow, B., Lin, X., Blackberry, I., Haapala, I., & Gaffy, E. … Hendy, S. (2012). *Healthy ageing literature review 2012*. Retrieved from http://www.health.vic.gov.au/agedcare/maintaining/downloads/healthy_litreview.pdf

Rosenberg, D., Depp, C. A., Vahia, I. V, Reichstadt, J., Palmer, B. W., Kerr, J., … Jeste, D. V. (2010). *Exergames for subsyndromal depression in older adults: a pilot study of a novel intervention. The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry* (Vol. 18, pp. 221–226).

Schaefer, C., Enderes, T., Ritter, H., & Zitterbart, M. (2002). Subjective Quality Assessment for Multiplayer Real-time Games. *Proceedings of the 1st Workshop on Network and System Support for Games* (pp. 74–78). New York, NY, USA: ACM. 10.1145/566500.566511

Shim, N., Baecker, R., Birnholtz, J., & Moffatt, K. (2010). TableTalk Poker. *Proceedings of the International Academic Conference on the Future of Game Design and Technology - Futureplay '10* (p. 98). New York, New York, USA: ACM Press. 10.1145/1920778.1920792

Stinson, J. N., Jibb, L. A., Nguyen, C., Nathan, P. C., Maloney, A. M., Dupuis, L. L., & ... (2013). Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. *Journal of Medical Internet Research*, *15*(3), e51. doi:10.2196/jmir.2350 PMID:23475457

Stress and Relationships. (2013). Retrieved from http://www.lifepositive.com/mind/psychology/stress/social-anxiety.asp

Strong, D., Tulu, B., Agu, E., & He, S. Q. (2014). Design of the Feedback Engine for a Diabetes Self-care Smartphone App. *Proceedings of Americas Conference on Information Systems - AMCIS 2014*. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1118&context=amcis2014

Tabak, M., Marin-Perianu, R. S., & Hermens, H. (2012). A serious game for COPD patients to perform physiotherapeutic exercises. In Nationale Longdagen (pp. 56–57). Utrecht, Netherlands: Book of abstracts. Netherlands Respiratory Society.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tamura, T., Yonemitsu, S., Itoh, A., Oikawa, D., Kawakami, A., Higashi, Y., & ... (2004). Is an Entertainment Robot Useful in the Care of Elderly People With Severe Dementia? *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *59*(1), M83–M85. doi:10.1093/gerona/59.1.M83 PMID:14718491

Taylor, A. H., Cable, N. T., Faulkner, G., Hillsdon, M., Narici, M., & Van Der Bij, A. K. (2004). Physical activity and older adults: A review of health benefits and the effectiveness of interventions. *Journal of Sports Sciences*, *22*(8), 703–725. doi:10.1080/02640410410001712421 PMID:15370483

Teixeira, P. J., Carraca, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: A systematic review. *The International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), 78. doi:10.1186/1479-5868-9-78 PMID:22726453

Theng, Y.-L., & Dahlan, A. Bin, Akmal, M. L., & Myint, T. Z. (2009). An Exploratory Study on Senior Citizens' Perceptions of the Nintendo Wii: The Case of Singapore. *Proceedings of the 3rd International Convention on Rehabilitation Engineering & Assistive Technology* (pp. 10:1–10:5). ACM. doi:10.1145/1592700.1592712

Thorsteinsen, K., Vittersø, J., & Svendsen, G. B. (2014). Increasing physical activity efficiently: An experimental pilot study of a website and mobile phone intervention. *International Journal of Telemedicine and Applications*, *746232*. doi:10.1155/2014/746232 PMID:24963290

Together for Health: A Strategic Approach for the EU 2008-2013 [White Paper]. (2007). Retrieved from http://ec.europa.eu/health/ph_overview/Documents/strategy_wp_en.pdf

Toscos, T., Faber, A., Connelly, K., & Upoma, A. M. (2008). Encouraging physical activity in teens can technology help reduce barriers to physical activity in adolescent girls? *Proceedings of the 2nd International Conference on Pervasive Computing Technologies for Healthcare 2008, PervasiveHealth* (pp. 218–221). 10.1109/PCTHEALTH.2008.4571073

United Nations - Department of Economic and Social Affairs Population Division. (2013). *World Population Ageing 2013* (p. 95). Retrieved from http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2013.pdf

Wada, K., Shibata, T., Saito, T., & Tanie, K. (2004). Effects of robot-assisted activity for elderly people and nurses at a day service center. *Proceedings of the IEEE*, *92*(11), 1780–1788. doi:10.1109/JPROC.2004.835378

Webster, D., & Celik, O. (2014). Systematic review of Kinect applications in elderly care and stroke rehabilitation. *Journal of Neuroengineering and Rehabilitation*, *11*(1), 108. doi:10.1186/1743-0003-11-108 PMID:24996956

Wieching, R., Kaartinen, N., DeRosario, H., Baldus, H., Eichberg, S., Drobics, M., & Delbaere, K. (2012). A new approach for personalized fall risk prediction & pre-vention: tailored exercises, unobtrusive sensing & advanced reasoning. *Journal of Aging and Physical Activity*, *20*, S120–S120.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Wiklund Axelsson, S., Melander Wikman, A., Näslund, A., & Nyberg, L. (2013). Older people's health-related ICT-use in Sweden. *Gerontechnology (Valkenswaard)*, *12*(1), 36–43. doi:10.4017/gt.2013.12.1.010.00

Williams, M. A., Soiza, R. L., Jenkinson, A. M., & Stewart, A. (2010). EXercising with Computers in Later Life (EXCELL) - pilot and feasibility study of the acceptability of the Nintendo® WiiFit in community-dwelling fallers. *BMC Research Notes*, *3*(1), 238. doi:10.1186/1756-0500-3-238 PMID:20831836

Wollersheim, D., & Merkes, M. (2010). Physical and psychosocial effects of Wii video game use among older women. *International Journal of Emerging Technologies and Society*, *8*(2), 85–98. Retrieved from http://www.helmholtz-muenchen.de/fileadmin/JOIN/PDF/Physical_and_Psychosocial_Effects_of_Wii_Video_Game_Use_among_Older_Women.pdf

Wüest, S., Borghese, N. A., Pirovano, M., Mainetti, R., van de Langenberg, R., & de Bruin, E. D. (2014). Usability and Effects of an Exergame-Based Balance Training Program. *Games for Health Journal*, *3*(2), 106–114. doi:10.1089/g4h.2013.0093 PMID:24804155

Wüest, S., Langenberg, R., & Bruin, E. D. (2013). Design considerations for a theory-driven exergame-based rehabilitation program to improve walking of persons with stroke. *European Review of Aging and Physical Activity*, 1–11. doi:10.100711556-013-0136-6 PMID:25309631

Yogyakarta Declaration on Ageing and Health 2012. (2012) (p. 3). Yogyakarta. Retrieved from http://kp.one.un.org/content/uploads/2012/09/Yogyakarta-Declaration-FINAL.pdf

Zelinski, E. M., & Reyes, R. (2009). Cognitive benefits of computer games for older adults. *Gerontechnology (Valkenswaard)*, *8*(4). doi:10.4017/gt.2009.08.04.004.00 PMID:25126043

ADDITIONAL READING

Anderson-Hanley, C., Snyder, A. L., Nimon, J. P., & Arciero, P. J. (2011). Social facilitation in virtual reality-enhanced exercise: Competitiveness moderates exercise effort of older adults. *Clinical Interventions in Aging*, *6*, 275–280. doi:10.2147/CIA.S25337 PMID:22087067

Baranowski, T., Baranowski, J., O'Connor, T., Lu, A. S., & Thompson, D. (2012). Is Enhanced Physical Activity Possible Using Active Videogames? *Games for Health Journal*, *1*(3), 228–232. doi:10.1089/g4h.2012.0018 PMID:24416640

Berkovsky, S., Freyne, J., & Coombe, M. (2010). Physical activity motivating games. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6206, pp. 283–284). LNAI.

Boulos, M. N. K., & Yang, S. P. (2013). Exergames for health and fitness: The roles of GPS and geosocial apps. *International Journal of Health Geographics*, *12*(1), 18. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3657542&tool=pmcentrez&rendertype=abstract doi:10.1186/1476-072X-12-18 PMID:23561306

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Brauner, P., Calero Valdez, A., Schroeder, U., & Ziefle, M. (2013). Increase Physical Fitness and Create Health Awareness through Exergames and Gamification. In A. Holzinger, M. Ziefle, M. Hitz, & M. Debevc (Eds.), *Human Factors in Computing and Informatics SE - 22* (Vol. 7946, pp. 349–362). Springer Berlin Heidelberg; doi:10.1007/978-3-642-39062-3_22

Chamberlin, B., Maloney, A., Gallagher, R. R., & Garza, M. L. (2013). Active Video Games: Potential for Increased Activity, Suggestions for Use, and Guidelines for Implementation. In S. Arnab, I. Dunwell, & K. Debattista (Eds.), *Serious Games for Healthcare: Applications and Implications* (pp. 191–212). Hershey, PA, USA: IGI Global; doi:10.4018/978-1-4666-1903-6.ch009

Drobics, M., & Smith, S. (2014). Game-Based IT Solutions for Active and Healthy Aging. In S. Göbel & J. Wiemeyer (Eds.), *Games for Training, Education, Health and Sports* (Vol. 8395, pp. 8–11). Springer International Publishing; doi:10.1007/978-3-319-05972-3_2

Garcia, J. A., Felix Navarro, K., Schoene, D., Smith, S. T., & Pisan, Y. (2012). Exergames for the elderly: Towards an embedded Kinect-based clinical test of falls risk. *Studies in Health Technology and Informatics*, *178*, 51–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22797019 PMID:22797019

Garcia Marin, J., Felix Navarro, K., & Lawrence, E. (2011). Serious Games to Improve the Physical Health of the Elderly: A Categorization Scheme. *Proceedings of CENTRIC 2011, The Fourth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services* (pp. 64–71). Retrieved from http://www.thinkmind.org/index.php?view=article&articleid=centric_2011_3_20_30056

Graves, L. E. F., Ridgers, N. D., Williams, K., Stratton, G., Atkinson, G., & Cable, N. T. (2010). The physiological cost and enjoyment of Wii Fit in adolescents, young adults, and older adults. *Journal of Physical Activity & Health*, *7*(3), 393–401. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20551497 PMID:20551497

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. *Proceedings of System Sciences (HICSS), 2014 47th Hawaii International Conference* (pp. 3025–3034). doi:10.1109/HICSS.2014.377

Kayama, H., Okamoto, K., Nishiguchi, S., Yamada, M., Kuroda, T., & Aoyama, T. (2014). Effect of a Kinect-based exercise game on improving executive cognitive performance in community-dwelling elderly: Case control study. *Journal of Medical Internet Research*, *16*(2), e61. doi:10.2196/jmir.3108 PMID:24565806

Korn, O., Brach, M., Hauer, K., & Unkauf, S. (2013). Exergames for Elderly Persons: Physical Exercise Software Based on Motion Tracking within the Framework of Ambient Assisted Living. In K. Bredl & W. Bösche (Eds.), *Serious Games and Virtual Worlds in Education, Professional Development, and Healthcare* (pp. 258–268). Hershey, PA, USA: IGI Global; doi:10.4018/978-1-4666-3673-6.ch016

Marques, A. I., Rosa, M. J., Amorim, M., Soares, P., Oliveira-Tavares, A., Santos, R., & ... (2012). Study protocol: Using the Q-STEPS to assess and improve the quality of physical activity programmes for the elderly. *BMC Research Notes*, *5*(1), 171. doi:10.1186/1756-0500-5-171 PMID:22471954

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Marques, A. I., Rosa, M. J., Soares, P., Santos, R., Mota, J., & Carvalho, J. (2011). Evaluation of physical activity programmes for elderly people - a descriptive study using the EFQM' criteria. *BMC Public Health*, *11*, 123. PMID:21338497

Marques, A. I., Soares, P., Soares-Miranda, L., Moreira, C., Oliveira-Tavares, A., Clara-Santos, P., & ... (2011). Evaluation of physical activity programmes for the elderly - exploring the lessons from other sectors and examining the general characteristics of the programmes. *BMC Research Notes*, *4*(1), 368. doi:10.1186/1756-0500-4-368 PMID:21943398

McKeon, P. O., & Hertel, J. (2008). Systematic Review of Postural Control and Lateral Ankle Instability, Part II: Is Balance Training Clinically Effective. *Journal of Athletic Training*, *43*(3), 305–315. doi:10.4085/1062-6050-43.3.305 PMID:18523567

Nap, H. H., & Diaz-Orueta, U. (2013). Rehabilitation Gaming. In S. Arnab, I. Dunwell, & K. Debattista (Eds.), *Serious Games for Healthcare: Applications and Implications* (pp. 50–75). Hershey, PA, USA: IGI Global; doi:10.4018/978-1-4666-1903-6.ch003

Nap, H. H., & Diaz-Orueta, U. (2014). Rehabilitation Gaming. In J. Bishop (Ed.), *Gamification for Human Factors Integration: Social, Education, and Psychological Issues* (pp. 122–147). Hershey, PA, USA: IGI Global; doi:10.4018/978-1-4666-5071-8.ch008

Némery, A., & Brangier, E. (2014). Set of Guidelines for Persuasive Interfaces: Organization and Validation of the Criteria. *Journal of Usability Studies*, *9*(3), 105–128. Retrieved from http://www.usabilityprofessionals.org/upa_publications/jus/2014may/JUS_Nemery_May_2014.pdf

Paredes, H., Cassola, F., Morgado, L., de Carvalho, F., Ala, S., & Cardoso, F. … Martins, P. (2014). Exploring the Usage of 3D Virtual Worlds and Kinect Interaction in Exergames with Elderly. In K. Miesenberger, D. Fels, D. Archambault, P. Peňáz, & W. Zagler (Eds.), Computers Helping People with Special Needs (Vol. 8547, pp. 297–300). Springer International Publishing. doi:10.1007/978-3-319-08596-8_47

Pasco, D., Bossard, C., Buche, C., & Kermarrec, G. (2010). Using Exergames to Promote Physical Activity : A Literature Review. *Sport Science Review*, *XIX*, 77–93. doi:10.2478/v10237-011-0048-1

Pereira, P., Duarte, E., Rebelo, F., & Noriega, P. (2014). A Review of Gamification for Health-Related Contexts. In A. Marcus (Ed.), *Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments SE - 70* (Vol. 8518, pp. 742–753). Springer International Publishing; doi:10.1007/978-3-319-07626-3_70

Pichierri, G., Wolf, P., Murer, K., & de Bruin, E. D. (2011). Cognitive and cognitive-motor interventions affecting physical functioning: A systematic review. *BMC Geriatrics*, *11*(1), 29. doi:10.1186/1471-2318-11-29 PMID:21651800

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology*, *14*(2), 154–166. doi:10.1037/a0019440

Sawyer, B. (2007). Serious games: Broadening games impact beyond entertainment. *Computer Graphics Forum*, *26*(3), xviii. doi:10.1111/j.1467-8659.2007.01044.x

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Studenski, S., Perera, S., Hile, E., Keller, V., Spadola-Bogard, J., & Garcia, J. (2010). Interactive video dance games for healthy older adults. *The Journal of Nutrition, Health & Aging*, *14*(10), 850–852. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21125204 doi:10.100712603-010-0119-5 PMID:21125204

Suárez, H., Suárez, A., & Lavinsky, L. (2006). Postural adaptation in elderly patients with instability and risk of falling after balance training using a virtual-reality system. *The International Tinnitus Journal*, *12*, 41–44. PMID:17147038

Wattanasoontorn, V., Boada, I., García, R., & Sbert, M. (2013). Serious games for health. *Entertainment Computing*, *4*(4), 231–247. doi:10.1016/j.entcom.2013.09.002

White, H., McConnell, E., Clipp, E., Branch, L. G., Sloane, R., Pieper, C., & Box, T. L. (2002). A randomized controlled trial of the psychosocial impact of providing internet training and access to older adults. *Aging & Mental Health*, *6*(3), 213–221. doi:10.1080/13607860220142422 PMID:12217089

This research was previously published in the Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice edited by Bengisu Tulu, Daniel Novák, and Håvar Brendryen; pages 245-276; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 81
Exploring the Role of Service Quality and Knowledge for Mobile Health Services

Nabila Nisha
North South University, Bangladesh

Mehree Iqbal
North South University, Bangladesh

Afrin Rifat
North South University, Bangladesh

Sherina Idrish
North South University, Bangladesh

ABSTRACT

The use of mobile technology-based services has made healthcare delivery more accessible and affordable in recent times. In fact, mobile health services today act as an effective means of providing healthcare knowledge to users directly from providers. However, the cynical behavior of users regarding this medium of healthcare services often encircles around the quality of such services. The aim of this paper is to examine the role of service quality and knowledge among other underlying factors that can influence future use intentions of m-Health services in the context of Bangladesh. The conceptual model of the study identifies that certain aspects of service qualities like reliability, privacy, responsiveness, empathy and information quality along with facilitating conditions, effort expectancy, performance expectancy and social influence plays an important role in capturing users' overall perceptions of mobile health services. Finally, the study highlights managerial implications, future research directions and limitations from the Bangladesh perspective.

DOI: 10.4018/978-1-7998-1204-3.ch081

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Information and communication technology (ICT) offers great potential for improving the quality of services provided along with the efficiency and effectiveness in health care sectors. In recent times, electronic health (e-Health) services truly serves as a tool with a huge potential for health care organizations to deliver quality and cost-effective care to geographically dispersed populations through the Internet (Jung, 2008}. However, a branch of e-Health services - mobile health (m-Health) has been constantly expanding over the last decade.

The m-Health services can be broadly defined as the use of mobile computing and communication technologies in health care and public health. This service application generally provides patient monitoring, sends text messages reminding patients to take needed medications and offers suggestions for maintaining health while pregnant, even in war-ravaged places. Current evidence suggests that the use of mobile technology can improve diagnosis and compliance with treatment guidelines, as well as patient information and can increase administrative efficiency (Rashidee, 2013}. In a broader sense, m-Health is not just improving health status rather it is the use of mobile technology to address health care challenges such as access, quality, affordability, behavioral norms, skill development in communication, supply management, information management and financial transactions through the exchange of information (Sultana, 2014}.

The m-Health services are a great initiative in health care sectors, since there are a number of patients who possess less knowledge and understanding of personal health problems but cannot afford time or money to visit doctors or medical centres on a regular basis (Nisha et al., 2015}. However, the benefits of m-Health services are often outweighed by the fact that many users are still skeptical of the quality of services that can be provided through such mediums. As such, it becomes imperative to examine the role of service quality and knowledge among other factors that can influence the acceptance, use, future prospect and necessity of m-Health services from the perspective of a developing country.

In the context of developing countries, technology may be well-perceived but when the content is sensitive like healthcare provisions, acceptance of the technology often depends upon the quality of the services and individual knowledge related to the service, among other factors. Therefore, the unified theory of acceptance and use of technology (UTAUT) model has been used for this paper. This study employed proposed constructs of system quality (system reliability, system efficiency, system privacy), information quality, interaction quality (responsiveness, assurance, empathy) and healthcare knowledge, to examine the factors that can influence users' intention to use m-Health services in Bangladesh. Moreover, this study has both theoretical and managerial implications. Theoretically, drawing upon relevant literature, this study aims to provide a model that is capable of understanding the determinants behind the future adoption of m-Health services among the people of Bangladesh. From a managerial perspective, the findings of this research should provide further insights into understanding and managing potential m-Health users, particularly hailing from the developing countries. This study can also assist various public and private hospitals and various telecommunication networks to consider the idea of providing m-Health services to the people of Bangladesh.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

LITERATURE REVIEW

This study argues that service quality and individual knowledge of health care plays a deterministic role in influencing the future use of m-Health services in the context of Bangladesh. As such, the current practice of m-Health services and its implications in Bangladesh, followed by the research platform of the study has been discussed to determine the gaps for the study.

A Budding Ground for Mobile Health Initiatives

Bangladesh is experiencing significant advances and development in health care sectors in recent years, since the government has started a new era in the health sector by introducing ICT for health service delivery. The government has developed a Health Management Information System (MIS) department under the Directorate General of Health Services (DGHS). The purpose of this department is to ensure the best use of ICT to build and maintain nationwide health information system of Bangladesh (DGHS, 2014}. This works as the backbone of the e-Health service network of Bangladesh and m-Health service is rendered using the national wide mobile phone network. The health professionals provide basic health advices and initial diagnosis when the service recipients' contacts through specialized 24/7 call centers. It has established free tele-consultation with government doctors, free SMS services for patient management and communication with staff, particularly for the needy people of the country. These cell phone numbers are circulated among the surrounding community. As such, people residing in the rural areas can contact with the health professionals through this network.

Additionally, web-camera has been installed in each sub-district, district, medical college and post-graduate institute hospitals in Bangladesh for providing telemedicine services using video conferencing platform. In recent times, there are few m-Health programmes that are operational in Bangladesh like m-Care, Aponjon and such others. These provide information and advice to pregnant mothers regarding health and nutrition related matters (Sultana, 2014}. In addition, the Mobile Alliance for Maternal Action (MAMA} Bangladesh program from D.Net with assistance from USAID and Johnson & Johnson provides voice messages to pregnant women to support the antennal, natal and post-natal information in Bengali (the native language of Bangladesh) (Reza, 2012}. Many mobile phone companies are also providing medical advice and prescriptions to millions of callers in their networks at nominal fees, alleviating some of the access and timeliness related challenges in the provision of healthcare services in Bangladesh (Rashidee, 2013}.

However, m-health service does not yet have a wide usage rate in Bangladesh. Across the entire country, m-Health services are being extensively adopted in the rural side of the country, which shows that the capital city of Dhaka is yet to adopt the use of such services. This paper, therefore, argues that factors like the service quality and necessary health care knowledge may influence the adoption and use of m-Health services among the dwellers of Dhaka city in Bangladesh.

Research Platform of the Study

Although a substantial section of the rural population uses m-Health applications today in Bangladesh, the urban people are still concerned about whether advice from an unknown hospital doctor should be considered. As such, evidence of quality of the m-Health services and use of individual knowledge of healthcare may bring about a change in the behavior of the urban people.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

One of the most important branches of information system research is to understand individual acceptance and use of information technology according to many literatures. The focus shifted towards technology acceptance and use, when the Technology Acceptance Model (TAM) was widely employed in many past studies (Davis, 1989; Davis et al., 1989}. Based on TAM synthesis of prior technology acceptance research, Venkatesh et al. (2003} developed the unified theory of acceptance and use of technology (UTAUT} model. The factors included in UTAUT (performance expectancy, effort expectancy, social influence and facilitating conditions) has been primarily used to predict the behavioral intention to use a technology and technology use in organizational contexts, moderated by individual difference variables like age, gender, experience and voluntariness (Venkatesh et al., 2012}.

In both organizational and non-organizational settings, UTAUT has repeatedly served as a baseline model to study a variety of technologies. However, given the number of technology devices, applications and services that are targeted at consumers in recent times, it became necessary to identify the factors that can influence consumer adoption and use of technologies (Stofega and Llamas, 2009}. This led to the introduction of the UTAUT2 model by Venkatesh et al. (2012}.

In UTAUT2, Venkatesh et al. (2012} adapted the four key constructs (i.e. performance expectancy, effort expectancy, social influence and facilitating conditions) that influence behavioral intention to use a technology and technology use from the original UTAUT model and customized it to fit the consumer context. Previous studies like Venkatesh et al. (2003} claim that the constructs of performance expectancy, effort expectancy and social influence determines the behavioral intention to use a technology, while behavioral intention and facilitating conditions influence the technology use in a particular context.

In addition, Venkatesh et al. (2012} further claimed that the addition of new constructs in a consumer context can contribute to the expansion of the theoretical horizons of the UTAUT model. Following the suit, this study has selected the original UTAUT and UTAUT2 model as a theoretical foundation to develop a proposed research model for the domain of healthcare from the consumers' perspective.

Proposed Constructs of Service Quality and Knowledge

Some recent studies like Sunyaev et al. (2014} and Hoque et al. (2015} argued that reliability, efficiency and privacy are the major dimensions of system quality that should be inspected in the context of sensitive mobile-based technologies like healthcare services. According to Hoque and Sorwar (2015}, users' perception of system reliability can be a major factor in capturing users' perceptions towards m-Health services in countries like Bangladesh, where mobile-based services are relatively new today. System efficiency, on the other hand, is the technical availability of the system, wherein accessibility, flexibility and response time are three major attributes. James et al. (2015} claimed that users with more positive impressions about accessibility, flexibility and response time are more willing to use m-Health services. The third dimension of system quality that has been explored in many studies is privacy. In the context of the adoption of technology-based services, privacy has always been considered as an important factor to gain reliance on the service platform as it protects personal information of the user (Ashrafi and Kuilboer, 2005; Guo et al., 2015}. Particularly, if it is customized mobile health services, these three factors of reliability, efficiency and privacy play a more significant role since health information requires confidentiality and maintenance of anonymity (Ren et al., 2010}.

Information quality is another additional construct that has been included in the proposed model of this study. This variable generally plays a critical role in building user satisfaction and customer loyalty for mobile healthcare services (Kumar et al., 2013}. According to Kahn et al. (2010}, information that is

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

provided to the patients over the mobile platform is mostly considered to be significant in such contexts. Therefore, it is very important that whatever information about healthcare services is provided should be complete, accurate and up-to-date (Fox, 2011}. Free et al. (2013} stated that more complete, accurate and up-to-date healthcare information will ultimately result into the enhancement of information quality. As such, this construct is deemed to be a crucial service quality that can influence the consumers to adapt the facilities of m-Health services.

Responsiveness, assurance and empathy are three other elements considered in this study and perceived under interaction quality of m-Health services. According to Yin et al. (2015}, interaction quality in the form of responsiveness, assurance and empathy affects user's loyalty indirectly and directly towards the use of mobile technology. Interaction quality is often vital in m-Health services, as people seek for customized attention with prompt consultancy in such matters (Nambisan, 2011}. Usually, the consumers seeking such services lack contextual emotional stability and hence timely responsiveness, accurate consultancy with emotional support and care tend to ease the conditional mindset at that context. Washington et al. (2015} claimed that relationship marketing has always played an important role in any service sector and so, empathy can lead to a strong relationship developed with the consumer base for such providers. Moreover, Xue et al. (2012} argued that by mentioning the name of the consultant providing the service in the delivered information and taking the accountability of credible consultancy, assurance can also be provided to the consumers. Even while delivering the service promptly over mobile devices, Huang et al. (2015} stated having the providers address the consumer with name and showing care with follow-up interactions, along with assurance and empathy the variable of responsiveness can be ensured as well.

Healthcare knowledge has been considered as an important factor in this study since healthcare is an area of which people are mostly sensitive. The probability of consumers seeking m-Health services is dependent on their prior knowledge on such. Gagnon et al. (2015} claimed that users lacking confidence about their relative knowledge related to healthcare are more likely to adopt this service. Townsend et al. (2015} further stated when knowledge is handy by other means, instead of seeking the service elsewhere consumers prefer using healthcare services over the mobile platform. On the other hand, Domínguez-Mayo et al. (2015} claimed that ensuring a basic healthcare knowledge among the citizen should be one of the major priorities of government for any developing economy. This is because such knowledge can facilitate the execution of m-Health services provided and ensures relatively quick relief, since consumers can relate their healthcare problems in a more proper way (Hasanain et al., 2015}.

RESEARCH MODEL AND HYPOTHESES

The purpose of this study is to examine the role of service quality and knowledge to determine the factors that can explain and predict users' intention to use m-Health services significantly in the context of Bangladesh. Along with the four key constructs (i.e. performance expectancy, effort expectancy, social influence and facilitating conditions) of the original model, four additional constructs, drawn from previous literature of m-Health services, has also been incorporated in this research model to make a significant theoretical contribution to the consumer context of the UTAUT model. The proposed research model used to address the influencing factors for healthcare technologies has been presented in Figure 1. In addition, all the variables hypothesized in this study and their likely relationships towards consumer acceptance and use of m-Health services in Bangladesh has been discussed next.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

System Quality

System quality is an overall measure of the information processing system itself. Even though system quality is sometimes perceived to be overlapped with factors that are closely related to service quality and ease of use, it has its own unique dimensions (Nelson et al., 2005}. The three most significant dimension of system quality are discussed below.

System Reliability

System reliability explains the accuracy of the technical functioning of the system and the truthfulness of service promises (Parasuraman et al., 2005}. According to Nelson et al. (2005}, it can be defined objectively as the technical availability of the system or the ability of the system to provide error free services. Thus, often consumers' perception of system reliability plays a vital role in defining consumers' adoption of m-Health services (Akter et al., 2010}. Particularly, for countries like Bangladesh where the concept of m-Health is relatively new, consumers' perception of reliability regarding m-Health is really questionable.

System Efficiency

From a consumer perspective, an efficient system is the one that is simple to use, structured properly, and requires minimum information to be input by individual (Parasuraman et al., 2005}. Therefore, how quickly consumers can reach the system when needed, how fast the system can be acclimatized to a variety of user needs and changing conditions and how quickly or timely the system responds to requests for information or action – all define system efficiency (Nelson et al., 2005}. This attribute usually has a strong influence on performance expectancy of m-Health services as they guide the perceptions of system quality and ultimately affect the usage intention of people.

System Privacy

Perceived privacy of consumer can be defined as the users' perception of protection against security threats and control of their personal health care information in an online environment (Lallmahmood, 2007}. It measures the degree to which the user believes the system is safe from intrusion and personal information is protected (Akter et al., 2010}. According to Angst and Agarwal (2009}, concern over privacy and security of personal health information is one of the reasons that held consumers from using m-Health services in developing countries like Bangladesh.

Hence, this study proposes the following hypotheses from the three characteristics of system quality:

H1A: System reliability positively influences performance expectancy of m-Health services.
H1B: System efficiency positively influences performance expectancy of m-Health services.
H1C: System privacy positively influences performance expectancy of m-Health services.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Information Quality

Lee and Chung (2009} suggest that information quality can significantly predict users' perception regarding the use of mobile based technology in sensitive areas like m-Health. Generally, the three most important indicators of information quality are completeness, accuracy and currency (Masrek et al., 2012}. Accuracy has been defined as the extent to which the information is correct, unambiguous, objective and meaningful or believable (Wand and Wang, 1996}. On the other hand, completeness refers to the extent to which all possible states relevant to the user population are available in the stored information, while currency can be defined as the degree to which the available information is up-to-date (Nelson et al., 2005}. As such, information quality can have a significantly positive effect on performance expectancy, which in turn can impact the behavioral intention of consumers towards m-Health. Thus, the study hypothesizes that:

H2: Information quality positively influences performance expectancy of m-Health services.

Interaction Quality

In this study, interaction quality has been perceived as a combination of three elements - responsiveness, assurance and empathy. Lin (2011} claims that interaction quality in the form of responsiveness, assurance and empathy affects users' loyalty indirectly and directly toward mobile internet technology. All these three factors are therefore discussed below.

Responsiveness

Responsiveness is the most important indicator of interaction quality (Lau et al., 2013}. It can be defined as the willingness to help users and to provide prompt services (Akter et al., 2010}. It can further be explained by the providers' ability to understand the users' problems and to provide precise services. Lau et al. (2013} found positive relationship between responsiveness and consumer loyalty. Based on this evidence, this study used responsiveness as a factor that will guide consumers regarding m-Health service adoption.

Assurance

Assurance measures knowledge, competency and courtesy of the provider that generate users' trust and confidence in the system (Akter et al., 2010}. It is the second most important factor of interaction quality that inspires consumer loyalty (Lau et al., 2013}. The service provider's ability to provide their promised and accurate healthcare services and their tendency of keeping commitments can only assure users regarding the adoption of mobile based technology. As past evidences from Lin (2011} and Lau et al. (2013} found positive relationship between assurance and consumers' acceptance of mobile technology, it has been perceived as an influencing factor for m-Health in this study.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Empathy

The final dimension of interaction quality is empathy that refers to the understanding ability of the users' needs and ability to provide personalized attention (Akter et al., 2010}. Consumers always appreciate a friendly and considerate environment. Hence, if the service provider shows genuine interest, enthusiasm and seriousness toward the users' healthcare needs, it will automatically lead consumers towards usage of m-Health services. In fact, Lau et al. (2013} found the highest correlation between empathy and customer satisfaction. This study thus proposes that empathy can positively influence performance expectancy, which can lead to the behavioral intention of consumers towards m-Health.

This study hypothesizes the following from the three characteristics of interaction quality:

H3A: Responsiveness positively influences performance expectancy of m-Health services.
H3B: Assurance positively influences performance expectancy of m-Health services.
H3C: Empathy positively influences performance expectancy of m-Health services.

Performance Expectancy

Burgess and Sargent (2007} and Wu et al. (2007} argue that the effect of performance expectancy is the most relevant factor for the adoption of internet-based technology. Performance expectancy generally depicts a users' view of the usefulness of adopting a technology (Venkatesh et al., 2003}. Sun et al. (2013} claim that in the context of m-Health services, the usefulness can only be captured by the extent to which it can help users to solve their health-related issues. If users believe that using m-Health services can help them to solve their problems, they are more likely to adopt this technology. Hence, the hypothesis is:

H4: Performance expectancy significantly affects individual intention to use m-Health services.

Effort Expectancy

Effort expectancy is considered to be directly related with the ease of using a particular technology (Phichitchaisopa and Naenna, 2013}. According to Venkatesh (1999}, all of these effort-oriented constructs act as more significant factors during the early stages of adopting a new technology. Several studies like Park et al. (2007}, Moores (2012} and Sun et al. (2013} claims that perceived ease of use or effort expectancy has considerable impacts on attitude towards the adoption of m-Health or any other healthcare related technology. As a result, the following hypothesis has been proposed:

H5: Effort expectancy significantly affects individual intention to use m-Health services.

Social Influence

Social influence refers to the degree to which an individual perceives that important others believe he or she should use the new system or technology (Venkatesh et al., 2003}. The idea behind social influence is that even though an individual may not be in favour of adopting a new technology, they intend to use it as he/she believes it will enhance his/her image among his/her family and peers (Venkatesh

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and Davis, 2000}. Researchers like Jung (2008} and Sun et al. (2013} empirically showed that there is a significant positive relationship between social influence and adoption of m-Health technology. Thus, the proposed hypothesis is:

H6: Social influence significantly affects individual intention to use m-Health services.

Facilitating Conditions

According to Venkatesh et al. (2003}, facilitating conditions refer to the resources and technical infrastructure that a user believes exists to support the adoption of a particular technology. In other words, facilitating conditions indicates the prospective conditions that may restrain or facilitate adopting a technology (Sun et al., 2013}. Venkatesh et al. (2012} claims that a consumer with a lower level of facilitating conditions can have a lower intention to use a particular technology. Moreover, Boontarig et al. (2012}, Phichitchaisopa and Naenna (2013} and Sun et al. (2013} showed that there is a positive significant relationship between facilitating conditions and healthcare technologies. Based on these findings, this study hypothesizes that:

H7: Facilitating conditions significantly affects individual intention to use m-Health services.

HealthCare Knowledge

Users' knowledge regarding healthcare can act as an important predictor of m-Health adoption. This is relatively a less investigated construct and was developed to measure the users' extent of knowledge and understanding of personal health problems (Wilson and Lankton, 2004}. Given all the resources and necessary infrastructure, people may tend to use internet-based health technologies to gain information regarding a health issue and increase their knowledge. This indicates that users who feel that they have relatively little knowledge about healthcare are more likely to adopt healthcare technologies. However, when Wilson and Lankton (2004} investigated this construct against the backdrop of e-Health technology they found it insignificant as a predictor of behavioral intention. To focus on this relationship further, this study hypothesizes that:

H8: Health care knowledge significantly affects individual intention to use m-Health services.

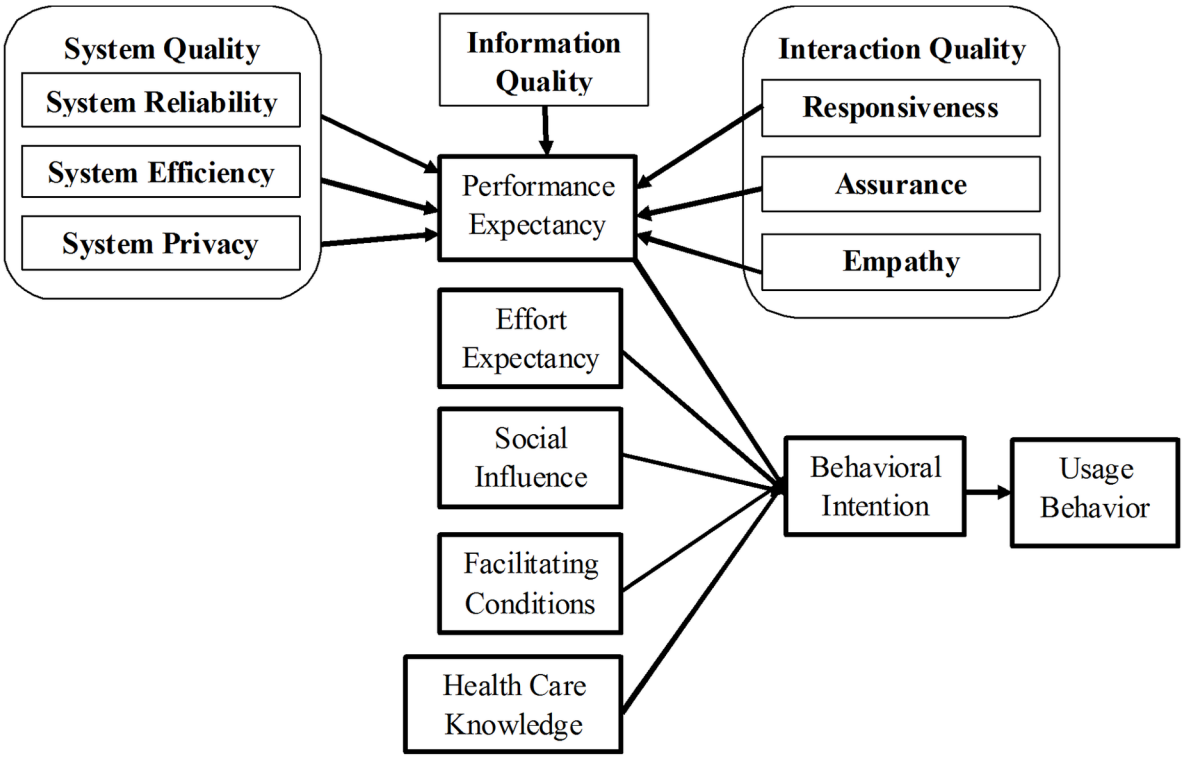
Behavioral Intention

Behavioural intention, which refers to the intention to use a system, is the major determinant of the actual behaviour. Researchers like Venkatesh and Zhang (2010} and Yu (2012} have repeatedly emphasized the strength of the construct of behavioural intention on usage behaviour. These past studies claim that individual behavior is predictable and can be influenced by individual intention that, in turn, can have a significant influence on technology usage. In the context of m-Health services, Jung (2008} and Sun et al. (2013} investigated and empirically proved that behavioural or adoption intention of the technology positively affects its usage. Following the lead, this study next hypothesizes that:

H9: Behavioral intention significantly affects individual behavior of using m-Health services.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. The proposed research structure



RESEARCH METHOD

Data Collection

The predominant existence of mobile health services in the rural areas has inadvertently raised the question as to why urban people still avoid the use of such services. In addition, the absence of city hospitals as the provider of m-Health services has made it imperative to examine the acceptance and use of m-Health services among the city dwellers of Bangladesh. As such, the population selected for this study only represented urban people who are currently exposed to the use of mobile phones and can thereby avail mobile health services in the future.

Data for this study was thus collected by conducting a survey through paper-based questionnaires on a sample of 1000 respondents in the capital city of Bangladesh – Dhaka. By using probability sampling and a stratified random sampling method, respondents were then selected for the sample. The use of this particular sampling method allowed us to avoid biasness in data and provided equal opportunity for all city dwellers who can be the potential users of mobile health services in Bangladesh.

Respondents were mostly educated and had English proficiency and so they were asked to self-administer the English version of the questionnaire. In other cases, the participants self-administered the Bengali version of the questionnaire or the research assistants themselves filled in the questionnaire based on the respondents' verbal responses.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Measurement

All the items used to measure the research variables of the survey were adapted from previous studies on information and technological advancements, e-Health services and m-Health services - with minor changes in wording to tailor them to the context of Bangladesh. This ensured the content validity of the questionnaire used to assess each constructs depicted in Figure 1.

The quantitative survey contained 42 statements in order to evaluate the original constructs of the UTAUT model and the eight new constructs proposed for the model, as listed in Table 1. The scales for the UTAUT constructs (performance expectancy, effort expectancy, social influence, facilitating conditions and behavioral intention) were adapted from Jung (2008}, Venkatesh and Zhang (2010} and Sun et al. (2013}. The scales for system quality (reliability, efficiency and privacy), information quality and interaction quality (responsiveness, assurance and empathy) were drawn from Nelson et al. (2005}, Lallmahamood (2007}, Jung (2008} and Lau et al. (2013}. On the other hand, the items used to assess health care knowledge were solely based on Wilson and Lankton (2004}. All these items were measured using a five-point Likert scale, ranging from "strongly disagree" to "strongly agree".

After a four-week survey, 927 completed and usable responses were obtained from the structured questionnaires. Table 2 represents the demographic information of respondents in terms of gender and age.

DATA ANALYSIS AND RESULTS

The research model was tested using the structural equation modeling (SEM) facilitates of SmartPLS (version 2.0). The method of partial least squares (PLS) was mainly chosen to conduct this analysis since a number of interaction terms have been included in the research model and PLS is capable of testing these effects [Chin et al., 2003}. As such, the measurement model was examined first to assess the reliability and validity of the constructs and then, the structural model was analyzed to examine the relationships hypothesized in the research model.

Measurement Model

Results from the measurement model are presented in Table 3 and Table 4 and it includes information about the reliability, validity, correlations and factor loadings. The composite reliabilities of the constructs ranged between 0.731 and 1.000, which exceeds the 0.7 cut-off value as recommended by Nunnally and Bernstein [1994}. The average variance extracted (AVE) was greater than 0.5 in all cases and greater than each square correlations, which indicates that the model has both convergent validity and discriminant validity [Fornell and Larcker, 1981}. Moreover, the internal consistency reliabilities (ICRs) of multi-item scales modeled with reflective indicators was 0.75 or greater, suggesting adequate reliability. The pattern of loadings and cross-loadings also supported internal consistency and discriminant validity, with some exceptions: one item from each construct of information quality, empathy (interaction quality), social influence, facilitating conditions, healthcare knowledge, and two items from system reliability (system quality) and system efficiency (system quality) were deleted due to their low loadings and high cross-loadings.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Constructs and corresponding items

Constructs	Corresponding Items	Items Sources
System Reliability	[SR1} I believe m-Health services operate reliably [SR2} I believe m-Health services perform reliably [SR3} I believe the operation of m-Health services is dependable	Nelson et al. [2005}
System Efficiency	[SE1} I feel m-Health services make health information easy to access [SE2} I feel m-Health services provide health information in a timely fashion [SE3} I feel m-Health takes too long to respond to my healthcare problems	Nelson et al. [2005}
System Privacy	[SP1} I trust the ability of m-Health services to protect my privacy [SP2} I feel safe when I release health information to m-Health services [SP3} I feel matters on security has no influence in using m-Health services	Lallmahamood [2007}
Information Quality	[IQ1} I believe m-Health services provide complete information [IQ2} I believe m-Health services provide up-to-date information [IQ3} I believe m-Health services provide relevant information	Jung [2008}
Responsiveness	[R1} I think m-Health provides precise healthcare services [R2} I think m-Health has the ability to understand my healthcare needs [R3} I think m-Health services are helpful	Lau et al. [2013}
Assurance	[As1} I think m-Health provides healthcare services as promised [As2} I think m-Health provides accurate healthcare services [As3} I think m-Health services will keep their commitments	Lau et al. [2013}
Empathy	[E1} I think m-Health has the knowledge to solve my healthcare problems [E2} I think m-Health has the enthusiasm to understand my healthcare needs [E3} I think m-Health services will put my healthcare needs in the first place	Lau et al. [2013}
Performance Expectancy	[PE1} Using m-Health services will improve my life quality [PE2} Using m-Health services will make my life more convenient [PE3} Using m-Health services will make me more effective in my life [PE4} Overall, I find m-Health services to be useful in my life	Sun et al. [2013}
Effort Expectancy	[EE1} Learning to use m-Health services is easy for me [EE2} Becoming skillful at using m-Health services is easy for me [EE3} Interaction with m-Health services is easy for me [EE4} Overall, I think m-Health services are easy to use	Sun et al. [2013}
Social Influence	[SI1} People who are important to me think that I should use m-Health services [SI2} People who are familiar with me think that I should use m-Health services [SI3} People who influence me think that I should use m-Health services [SI4} Most people surrounding me use m-Health services	Jung [2008}, Sun et al. [2013}
Facilitating Conditions	[FC1} Using m-Health services suits my living environment [FC2} Using m-Health services fits into my working style [FC3} Using m-Health services is compatible with my life [FC4} Help is available if I have problems in using m-Health services	Venkatesh & Zhang [2010}
Health Care Knowledge	[HCK1} I am very knowledgeable regarding care for my health problems [HCK2} I understand my health problems and how to care for them	Wilson & Lankton [2004}
Behavioral Intention	[BI1} I prefer to use m-Health services [BI2} I intend to use m-Health services [BI3} I plan to use m-Health services	Sun et al. [2013}

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Demographic profile of respondents

Demographics	Frequency	Percentage (%)
Gender		
Male	546	58.9
Female	381	41.1
Age		
20 or below	158	17.0
21 – 30	476	51.3
31 – 40	145	15.6
41 – 50	94	10.1
Above 50	54	5.8

Structural Model

The path coefficients and significance levels in the structural model are presented in Figure 2. Among the three constructs of system quality, system reliability (0.240, $p<0.05$) and system privacy (0.120, $p<0.05$) showed significant and positive path towards performance expectancy, while system efficiency (0.084, $p>0.05$) reported an insignificant path. The factor of information quality (0.100, $p<0.05$) also displayed a significant and positive path towards performance expectancy. On the other hand, responsiveness (0.135, $p<0.05$) and empathy (0.072, $p<0.05$) under interaction quality had a significant and positive path towards performance expectancy, but assurance (0.070, $p>0.05$) showed an insignificant path.

Among other factors, facilitating conditions (0.235, $p<0.05$), effort expectancy (0.169, $p<0.05$), performance expectancy (0.155, $p<0.05$) and social influence (0.134, $p<0.05$) showed significant and positive paths to the behavioral intention of using mobile health services, in their order of influencing strength. However, the construct of healthcare knowledge (-0.024, $p>0.05$) reported an insignificant path towards the individual behavior of using mobile health services. Therefore, all hypotheses (except H1B, H3B and H8) dealing with behavioral intention to use mobile health services are supported. Subsequently, the hypothesized relationship between behavioral intention and usage (0.251, $p<0.05$) is found to be statistically significant, thereby supporting hypothesis H9.

FINDINGS AND IMPLICATIONS

This study has revealed quite a few understandings on the back stage possibilities of the potential usage of mobile health services among the urban population of Bangladesh. Some of the constructs representing service quality positively influences the performance expectancy while factors like facilitating conditions, effort expectancy, performance expectancy and social influence directly stimulates the usage of m-Health services. However, healthcare knowledge has been identified as an insignificant factor of behavioral intention in the context of Bangladesh. The inclined reasons behind such findings are quite obvious and most of such are shaped by the local culture and infrastructure of the country.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Factor loadings, composite reliability and AVEs

Constructs	Items	Factor Loadings	Composite Reliability	AVE
System Reliability	SR1	1.000	1.000	1.000
System Efficiency	SE1	1.000	1.000	1.000
System Privacy	SP1	0.755	0.789	0.555
	SP2	0.731		
	SP3	0.749		
Information Quality	IQ1	0.908	0.887	0.797
	IQ2	0.876		
Responsiveness	R1	0.801	0.844	0.643
	R2	0.858		
	R3	0.743		
Assurance	As1	0.825	0.881	0.711
	As2	0.849		
	As3	0.855		
Empathy	E2	0.833	0.875	0.778
	E3	0.928		
Performance Expectancy	PE1	0.807	0.889	0.668
	PE2	0.847		
	PE3	0.850		
	PE4	0.761		
Effort Expectancy	EE1	0.822	0.895	0.680
	EE2	0.834		
	EE3	0.840		
	EE4	0.802		
Social Influence	SI1	0.867	0.891	0.732
	SI2	0.876		
	SI3	0.823		
Facilitating Conditions	FC1	0.776	0.864	0.680
	FC2	0.877		
	FC3	0.818		
Health Care Knowledge	HCK2	1.000	1.000	1.000
Behavioral Intention	BI1	0.868	0.891	0.731
	BI2	0.873		
	BI3	0.823		

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

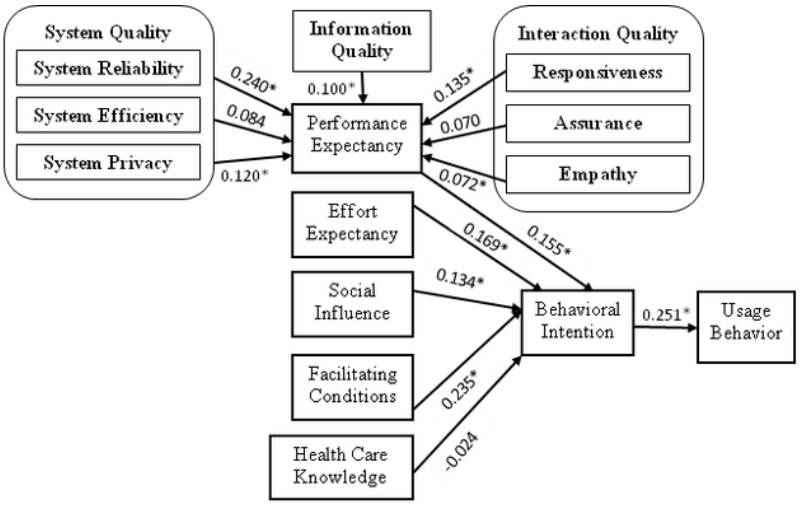
Table 4. Measurement model estimations

	ICRs	As	BI	E	EE	FC	HCK	IQ	PE	R	SE	SI	SP	SR
As	0.8	**0.8**												
BI	0.8	0.3	**0.9**											
E	0.8	0.1	0.0	**0.9**										
EE	0.8	0.3	0.5	0.0	**0.8**									
FC	0.8	0.3	0.5	0.0	0.4	**0.8**								
HCK	1.0	0.0	0.1	0.0	0.1	0.1	**Single-Item Construct**							
IQ	0.8	0.3	0.5	0.0	0.4	0.8	0.1	**0.9**						
PE	0.8	0.3	0.4	0.0	0.6	0.4	0.1	0.4	**0.8**					
R	0.8	0.2	0.4	0.1	0.3	0.3	0.1	0.3	0.3	**0.8**				
SE	1.0	0.8	0.2	0.1	0.3	0.3	0.1	0.3	0.3	0.2	**Single-Item Construct**			
SI	0.8	0.3	0.4	0.1	0.4	0.3	0.2	0.3	0.3	0.2	0.2	**0.9**		
SP	0.8	0.3	0.5	0.1	0.4	0.7	0.2	0.7	0.4	0.3	0.3	0.7	**0.7**	
SR	1.0	0.2	0.4	0.0	0.4	0.3	0.1	0.3	0.4	0.3	0.2	0.2	0.3	**Single-Item Construct**

Notes:

1. As (Assurance); BI (Behavioral Intention); E (Empathy); EE (Effort Expectancy); FC (Facilitating Conditions); HCK (Health Care Knowledge); IQ (Information Quality); PE (Performance Expectancy); R (Responsiveness); SE (System Efficiency); SI (Social Influence); SP (System Privacy); SR (System Reliability).
2. Diagonal elements represent the AVEs, while off-diagonal elements represent the square correlations.

Figure 2. Results of structural equation modeling



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Facilitating conditions is found as the strongest direct determinant in influencing respondents' behavioral intention of m-Health services. This finding is consistent to Boontarig et al. [2012}, Phichitchaisopa and Naenna [2013}, Sun et al. [2013} and Nisha et al. [2015}. The penetration of mobile phones is a comparatively new concept in Bangladesh and particularly, when such devices are being used as a platform of commuting other errands such as receiving instant health care advices, ensuring facilitating conditions has a strong role to play. In perspective of Bangladesh, facilitating conditions generally tend to be wide availability of computers, strong mobile networks and fast internet services. Therefore, the first business implication is that, beyond offering ease-of-use and useful m-Health services, providers might emphasize on the compatibility between the offered services and the working/living styles of their target customers. That is, putting efforts in designing suitable services and infrastructures to meet specific needs of different customer segments. Both government and private sector hospitals and clinics have already started showing a lot of interests and concerns in this regard. As a matter of fact, entities like Medinova Telemedicine, eClinic24 (Chakaria Project), AMCARE and health services for expatriates in Singapore by TRCL, Breast Cancer Finding via mobile by Amader Gram, JBFH Telemedicine, all leading telecom companies of Bangladesh and many others have already set their foot targeting the rural segments and working on developing the infrastructure challenges. However, city hospitals need to work more towards the development of such conditions for the city dwellers. Strong and correct comprehension on available facilitating conditions is important to ensure to build the right perception in the mind of the target market by showing repeatedly that all they need is a mobile phone of any type and a mobile connection mainly to take the benefits of mobile health care services.

The study drills that most of the current and potential users of m-Health services tend to judge the importance of the service based on the required effort expectancy for it. Respondents who expect less effort input in the service consuming process, especially at the adoption stage like that of m-Health services in Bangladesh, are more likely to show positive attitude in embracing the service. This finding is consistent with the evidence provided by Moores [2012}, Sun et al. [2013} and Nisha et al. [2015}. By making sure that the target market perceives the efforts needed to avail these services as an ease, companies can net a positive attitude towards these services. Communications of the providers, particularly of the city hospitals, should focus more on the actual handiness of such services. To do this, the software platform of the m-Health service applications should be made more compatible and user-friendly for the ease of use of the urban population of Bangladesh.

Performance expectancy is the next most important construct according to the research findings of this study. Sustainable qualities of m-Health service such as saving time, receiving immediate and accurate healthcare advices, etc. play an important role in shaping behavior in terms of adopting such new technologies. Jimison and Sher [2008}, Holden and Karsh [2010} and Sun et al. [2013} have also found such similar results in their studies. The underlying cause behind such result is the novelty of such technology in the culture of Bangladesh, where most of the errands are preferred to be served in a traditional manner from physical locations. Take away for marketers from this finding is to ensure a lot of confidence among the users by highlighting the value propositions of such service with their importance in their daily life through the marketing mix of the companies. Clinching benefits such as the quality of system and of information, high responsiveness with empathy, ease of using m-Health services and the relative conveniences (outcomes) that are attached with such usages should be used as the major selling ideas of their communication campaigns to build a strong positive comprehension in the response process of the target market. The providers must also ensure a lot of empathy and a high level of responsiveness while providing the accurate information through the call center personnel to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the users. They must keep in mind that one negative case in terms of accurate and quick response with empathy may lead to decreasing brand equity and hence, hospitals and clinics should train the front liners accordingly. Developing a sustainable system which ensures privacy and efficiency to support the clinching benefits of m-Health services is additionally important to encourage the consumption of such services among the city dwellers of Bangladesh.

Since the culture of Bangladesh is mainly based on collectivism, this is the root cause behind the impact of the factor of social influence on behavioral intention towards m-Health services. Being consistent with the findings of Jung [2008} and Sun et al. [2013}, which argued that subjective norm was the most influential factor for technologically related services, this study identified social influence as a powerful factor that can affect the intention to use m-Health services in Bangladesh. Further analysis revealed that respondents are significantly influenced by peer groups and interpersonal world-of-mouth, which is also consistent to the findings of Venkatesh and Davis [2000}. Taking the above together, the business implication for the m-Health service providers is to create a buzz to get the benefit of social capita. Semiotics such as catchy jingle and slogans with positive testimonials might also help the providers to bag a positive viral response from the market. Creative publicity and sales promotional activities using a lot of support media, particularly targeting the city dwellers may lead to a viral word of mouth. Social media such as Facebook and You Tube may also be exploited to get the viral going in urban markets.

Alternatively, the construct of healthcare knowledge has been noted to have an insignificant role to play in terms of shaping behavioral intention for m-Health services in Bangladesh. As per the Human Development Report by UNDP (United Nations Development Programme), the HDI (Human Development Index) of Bangladesh is 0.558 with a positive growth. This indicates that the healthcare knowledge of general citizens is below the line and most of the people, particularly in rural areas, are unaware of basic healthcare processes which make Bangladesh a potential for m-Health service providers. Take away for marketers will be to capitalize this dependency and to develop better ways to provide m-Health services in order to support the growth of the HDI index of Bangladesh. Also, the busy lifestyle of the city dwellers in Bangladesh often adds to this dependency of people on online or mobile health care services, which is why city hospitals should think about introducing such services in the urban areas as well.

Once city hospitals manage to introduce such services for the urban population, they can even think of establishing criterions like experience and original certifications of doctors before including them in the panel of the medical team that involves in the provision of m-Health services. This will be instrumental in reducing the concerns of the urban population regarding whether to take advice from unknown hospital doctors. Furthermore, it can accelerate the acceptance and use of m-Health services in Bangladesh if the urban population can trust the quality of the doctors and therefore, the quality of m-Health services in turn.

LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

This study has several limitations like other empirical past studies. First, the conclusions drawn from this study are based solely on the urban population, specifically people residing in Dhaka city of Bangladesh. Future research can explore the behavioral intention of people of other major cities or even the rural areas of the country. Second, a longitudinal study can be adopted in future works to examine and compare the research model in different time periods, thereby providing a better insight into the adoption of m-Health services in Bangladesh. Moreover, the model has been generalized and is only

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tested in a very specific condition of healthcare mobile services in a single country. The application of a similar conceptual model to study the context of other developing countries in South Asia like India and Pakistan can add further value. Also, panel data can be gathered regarding m-Health services from different developing countries and compared to do a more constructive analysis. Finally, other than the factors used in this study, there can be various other variables that can influence the acceptance and use of m-Health services. Further research considering different constructs can enhance the understanding of precise determinants for m-Health services in Bangladesh.

REFERENCES

Akter, S., D'Ambra, J., & Ray, P. (2010). User perceived service quality of m-Health services in developing countries. *Proceedings of the 18th European Conference on Information Systems: conference proceedings*, University of Pretoria. Retrieved from http://ro.uow.edu.au/cgi/viewcontent.cgi?article=4188&context=commpapers

Angst, M. C., & Agarwal, R. (2009). Adoption of electronic health records in the presence of privacy concerns: The Elaborate Likelihood Model and individual persuasion. *Management Information Systems Quarterly*, *33*(2), 339–370.

Ashrafi, N., & Kuilboer, J. (2005). Privacy Protection via Technology: Platform for Privacy Preferences (P3P). *International Journal of E-Business Research*, *1*(2), 56–69. doi:10.4018/jebr.2005040104

Boontarig, W., Chutimaskul, W., Chongsuphajaisiddhi, W., & Papasratorn, B. (2012). Factors influencing the Thai elderly intention to use smartphone for e-Health services. *Proceedings of the IEEE Symposium on Humanities, Science and Engineering Research* (pp. 242-246). 10.1109/SHUSER.2012.6268881

Burgess, L., & Sargent, J. (2007). Enhancing user acceptance of mandated mobile health information systems: The ePOC (electronic point-of-care project) experience. *Studies in Health Technology and Informatics*, *129*(2), 1088–1092. PMID:17911883

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo Simulation study and an Electronic-Mail Emotion/Adoption study. *Information Systems Research*, *14*(2), 189–217. doi:10.1287/isre.14.2.189.16018

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly*, *13*(3), 319–340. doi:10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003. doi:10.1287/mnsc.35.8.982

Directorate General of Health Services (DGHS). (2014). *Health information system & e-Health*. Retrieved from http://www.dghs.gov.bd/index.php/en/ehealth/our-ehealth-eservices/84-english root/ehealth-eservice/97-telimedicine-services-in-comunity-clinics

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Domínguez-Mayo, F. J., Escalona, M. J., Mejías, M., Aragón, G., García-García, J. A., Torres, J., & Enríquez, J. G. (2015). A Strategic Study about Quality Characteristics in e-Health Systems Based on a Systematic Literature Review. *TheScientificWorldJournal*, *2015*, 1–11. doi:10.1155/2015/863591 PMID:26146656

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *JMR, Journal of Marketing Research*, *18*(1), 39–50. doi:10.2307/3151312

Fox, S. (2011). *The social life of health information 2011*. Washington, DC: Pew Internet & American Life Project.

Free, C., Phillips, G., Watson, L., Galli, L., Felix, L., Edwards, P., & Haines, A. (2013). The effectiveness of mobile-health technologies to improve health care service delivery processes: A systematic review and meta-analysis. *PLoS Medicine*, *10*(1), e1001363. doi:10.1371/journal.pmed.1001363 PMID:23458994

Gagnon, M. P., Ngangue, P., Payne-Gagnon, J., & Desmartis, M. (2015). m-Health Adoption by Healthcare Professionals: A Systematic Review. *Journal of the American Medical Informatics Association*. PMID:26078410

Guo, X., Zhang, X., & Sun, Y. (2015). The privacy–personalization paradox in mHealth services acceptance of different age groups. *Electronic Commerce Research and Applications*. doi:10.1016/j.elerap.2015.11.001

Hasanain, R. A., Vallmuur, K., & Clark, M. (2015). Electronic Medical Record Systems in Saudi Arabia: Knowledge and Preferences of Healthcare Professionals. *Journal of Health Informatics in Developing Countries*, *9*(1).

Holden, R. J., & Karsh, B. T. (2010). The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics*, *43*(1), 159–172. doi:10.1016/j.jbi.2009.07.002 PMID:19615467

Hoque, M. R., Karim, M. R., & Amin, M. B. (2015). Factors Affecting the Adoption of mHealth Services among Young Citizen: A Structural Equation Modeling (SEM) Approach. *Asian Business Review*, *5*(2), 60–65.

Hoque, M. R., & Sorwar, G. (2015). Factors Influencing mHealth Acceptance among Elderly People in Bangladesh. *Proceedings of the Australasian Conference on Information Systems*, Australia.

Huang, E. Y., Lin, S. W., & Fan, Y. C. (2015). MS-QUAL: Mobile service quality measurement. *Electronic Commerce Research and Applications*, *14*(2), 126–142. doi:10.1016/j.elerap.2015.01.003

James, D. C., Harville, C. II, Whitehead, N., Stellefson, M., Dodani, S., & Sears, C. (2015). Willingness of African American women to participate in e-Health/m-Health research. *Telemedicine Journal and e-Health*. doi:10.1089/tmj.2015.0071 PMID:26313323

Jimison, H. B., & Sher, P. P. (2008). Consumer health informatics: Health information technology for consumers. *Journal of the American Society for Information Science*, *46*(10), 783–790. doi:10.1002/(SICI)1097-4571(199512)46:10<783::AID-ASI11>3.0.CO;2-L

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jung, M. (2008). From health to e-Health: Understanding citizens' acceptance of online health care, doctoral thesis, Luleå University of Technology, Sweden. Retrieved from http://epubl.ltu.se/1402-1544/2008/68/LTU-DT-0868-SE.pdf

Kahn, J. G., Yang, J. S., & Kahn, J. S. (2010). Mobile health needs and opportunities in developing countries. *Health Affairs*, *29*(2), 252–258. doi:10.1377/hlthaff.2009.0965 PMID:20348069

Kumar, S., Nilsen, W., Pavel, M., & Srivastava, M. (2013). Mobile health: Revolutionizing healthcare through trans-disciplinary research. *Computer*, *1*(1), 28–35. doi:10.1109/MC.2012.392

Lallmahamood, M. (2007). An examination of individual's perceived security and privacy of the Internet in Malaysia and the influence of this on their intention to use e-Commerce: Using an extension of the Technology Acceptance Model. *Journal of Internet Banking and Commerce*, *12*(3), 1–26.

Lau, M. M., Cheung, R., Lam, A. Y. C., & Chu, Y. T. (2013). Measuring service quality in the banking industry: A Hong Kong based study. *Contemporary Management Research*, *9*(3), 263–282. doi:10.7903/cmr.11060

Lee, K. C., & Chung, N. (2009). Understanding factors affecting trust in and satisfaction with mobile banking in Korea: A modified DeLone and McLean's model perspective. *Interacting with Computers*, *21*(5), 385–392. doi:10.1016/j.intcom.2009.06.004

Lin, H. (2011). An empirical investigation of mobile banking adoption: The effect of innovation attributes and knowledge-based trust. *International Journal of Information Management*, *31*(3), 252–260. doi:10.1016/j.ijinfomgt.2010.07.006

Lin, H. (2011). The effect of multi-channel service quality on mobile customer loyalty in an online-and-mobile retail context. *Service Industries Journal*, *32*(11), 1865–1882. doi:10.1080/02642069.2011.559541

Masrek, M. N., Uzir, N. A., & Khairuddin, I. I. (2012). Trust in mobile banking adoption in Malaysia: A conceptual framework. *Journal of Mobile Technologies*. *Knowledge in Society*, 2012, 1–12.

Moores, T. T. (2012). Towards an integrated model of IT acceptance in healthcare. *Decision Support Systems*, *53*(3), 507–516. doi:10.1016/j.dss.2012.04.014

Nambisan, P. (2011). Information seeking and social support in online health communities: Impact on patients' perceived empathy. *Journal of the American Medical Informatics Association*, *18*(3), 298–304. doi:10.1136/amiajnl-2010-000058 PMID:21486888

Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of Management Information Systems*, *21*(4), 199–235.

Nisha, N., Iqbal, M., Rifat, A., & Idrish, S. (2015). Mobile health services: A new paradigm for health care systems. *International Journal of Asian Business and Information Management*, *6*(1), 1–18.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory*. New York: McGraw-Hill.

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research*, *7*(3), 213–233. doi:10.1177/1094670504271156

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Park, J. K., Yang, S. J., & Lehto, X. (2007). Adoption of mobile technologies for Chinese consumers. *Journal of Electronic Commerce Research*, *8*(3), 196–206.

Phichitchaisopa, N., & Naenna, T. (2013). Factors affecting the adoption of healthcare information technology. *EXCLI Journal*, *12*, 413–436. PMID:26417235

Rashidee, A. H. (2013, June 30). Emerging mobile health in Bangladesh. *The Daily Star*. Retrieved from http://archive.thedailystar.net/beta2/news/emerging-mobile-Health-in-bangladesh/

Ren, Y., Pazzi, R. W. N., & Boukerche, A. (2010). Monitoring patients via a secure and mobile healthcare system. *Wireless Communications, IEEE*, *17*(1), 59–65. doi:10.1109/MWC.2010.5416351

Reza, P. R. (2012, December 26). Bangladesh: Mobile health service for expecting and new mothers, *Global Voice*. Retrieved from http://globalvoicesonline.org/2012/12/26/bangladesh-mobile-Health-service-for-expecting-and-new-mothers/

Stofega, W., & Llamas, R. T. (2009). *Worldwide Mobile Phone 2009-2013 Forecast Update. IDC Document Number 217209*. Framingham, MA: IDC.

Sultana, S. (2014, February 15). Taking care of health through mobile phones. *The Financial Express*, viewed 12 August 2014, <http://www.thefinancialexpress-bd.com/2014/02/15/18873>

Sun, Y., Wang, N., Guo, X., & Peng, Z. (2013). Understanding the acceptance of mobile health services: A comparison and integration of alternative models. *Journal of Electronic Commerce Research*, *14*(2), 183–200.

Sunyaev, A., Dehling, T., Taylor, P. L., & Mandl, K. D. (2014). Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*. doi:10.1136/amiajnl-2013-002605 PMID:25147247

Townsend, A., Leese, J., Adam, P., McDonald, M., Li, L. C., Kerr, S., & Backman, C. L. (2015). eHealth, Participatory Medicine, and Ethical Care: A Focus Group Study of Patients' and Health Care Providers' Use of Health-Related Internet Information. *Journal of Medical Internet Research*, *17*(6), e155. doi:10.2196/jmir.3792 PMID:26099267

Venkatesh, V. (1999). Creating favorable user perceptions: Exploring the role of intrinsic motivation. *Management Information Systems Quarterly*, *23*(2), 239–260. doi:10.2307/249753

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. doi:10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems Quarterly*, *27*(3), 425–478.

Venkatesh, V., Thong, J. Y. L., & Xin, X. (2012). Consumer acceptance and use of information technology: Extending the Unified Theory of Acceptance and Use of Technology. *Management Information Systems Quarterly*, *36*(1), 157–178.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Venkatesh, V., & Zhang, X. (2010). Unified theory of acceptance and use of technology: U.S. vs. China. *Journal of Global Information Technology Management*, *13*(1), 5–27. doi:10.1080/1097198X.2010.10856507

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality in ontological dimensions. *Communications of the ACM*, *39*(11), 86–96. doi:10.1145/240455.240479

Washington, G., Ward, J., & Kameka, M. (2015). Spare Me: Towards an Empathetic Tool for Helping Adolescents & Teenagers Cope with Sickle Cell. *Proceedings of the 2015 International Conference on Healthcare Informatics (ICHI)* (pp. 555-561). IEEE.

Wilson, E. V., & Lankton, N. K. (2004). Modeling patients' acceptance of provider-delivered E-health. *Journal of the American Medical Informatics Association*, *11*(4), 241–248. doi:10.1197/jamia.M1475 PMID:15064290

Wu, J., Wang, S., & Lin, L. (2007). Mobile computing acceptance factors in the healthcare industry: A structural equation model. *International Journal of Medical Informatics*, *76*(1), 66–77. doi:10.1016/j.ijmedinf.2006.06.006 PMID:16901749

Xue, L., Yen, C.C., Chang, L., Tai, B.C., Chan, H.C., Duh, H.B.L. & Choolani, M. (2012). Journeying Toward Female-focused m-Health Applications. In *Advances in Affective and Pleasurable Design* (p. 295).

Yin, S. Y., Huang, K. K., Shieh, J. I., Liu, Y. H., & Wu, H. H. (2015). Tele-health services evaluation: A combination of SERVQUAL model and importance-performance analysis. *Quality & Quantity*, 2015, 1–16.

Yu, C. (2012). Factors affecting individuals to adopt mobile banking: Empirical evidence from the UTAUT model. *Journal of Electronic Commerce Research*, *13*(2), 104–121.

This research was previously published in the International Journal of E-Business Research (IJEBR), 12(2); edited by Jeffrey Hsu; pages 45-64; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 82
Effect of Power and Phase Synchronization in Multi-Trial Speech Imagery

Sandhya Chengaiyan
SSN College of Engineering, India

Divya Balathayil
SSN College of Engineering, India

Kavitha Anandan
SSN College of Engineering, India

Christy Bobby Thomas
M.S.Ramaiah University of Applied Sciences, India

ABSTRACT

Speech imagery is one form of mental imagery which refers to the imagining of speaking a word to oneself silently in the mind without any articulation movement. In this work, electroencephalography (EEG) signals were acquired while speaking and during the imagining of speaking consonant-vowel-consonant (CVC) words in multiple trials of different time frames. Relative powers were computed for each EEG frequency band. It has been observed that relative power of alpha and theta bands was dominant. Phase Locking Value (PLV), a functional brain connectivity parameter has been estimated to understand the phase synchronicity between two brain regions. PLV results show that the left hemispheric frontal and temporal electrodes has maximum phase lock in alpha and theta band during speech and speech imagery process. The combination of brain connectivity estimators and signal processing techniques will thus be a reliable framework for understanding the nature of speech imagery signals captured through EEG.

DOI: 10.4018/978-1-7998-1204-3.ch082

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Mental Imagery is a quasi-perceptual experience which is induced in an absence of external stimuli (Perrone-Bertolotti et al., 2014). Mental Imagery occurs in any of the sense forms- Visual imagery (mind's eye), Auditory imagery (mind's ear), Motor imagery (mind's muscle movement), Tactile imagery (mind's touch). Speech imagery is one form of mental imagery which refers to imagining of speaking a word or sentence to oneself silently in the mind without any articulation movement (Tian & Poeppel, 2012). Speech imagery is also referred to as articulation imagery. The auditory and kinesthetic response induced in speech imagery involves both motor and perceptual systems (Kosslyn, 1994). The speech perception and production of speech must depend upon each other to ensure a proper communication as well as the content of messages being conveyed (Liberman & Mattingly, 1985). Children with Autism disorder have an excellent mind voice like other healthy controls. But Autism children find it difficult to express their thought due to their speech impairment. The concept of speech imagery can be helpful in decoding the thought for such children. Speech imagery can also contribute for speech preparation that has to convey from the decoded thought.

Speech processing theories have predicted that multiple brain regions participate in speech and language comprehension, which includes frontal and temporal lobes of left hemisphere. Broca's area found in the left inferior frontal gyrus and Wernicke's area located in the left posterior superior temporal gyrus are cortical areas specialized for production and comprehension, of human language respectively (Poeppel & Hickok, 2007; Price, 2011). Neuroscientists have followed conventional brain imaging techniques such as fMRI (Dogil et al., 2002) and electrophysiological measures such as EEG (Callan et al., 2000) to study the neuronal activation in these brain regions during speech production from healthy and speech impaired participants.

Electroencephalography (EEG), a non-invasive electrophysiological technique, is used in analyzing the neuronal excitations in different regions of the brain and to understand their functional cooperation while performing a task (Teplan, 2002). EEG signals contain useful information and this information can be analyzed in time domain, frequency domain and time-frequency domain. Transforms like Discrete Fourier Transform (DFT) and Fast Fourier Transform (FFT) can be used for converting a time domain signal into frequency domain in order to analyze the signal information in each frequency band over a wide range of frequencies related to the task. EEG signals have five sub-band frequencies such as Delta band (05.-3Hz), Theta band (4- 8Hz), Alpha band (8-12Hz), Beta band (13-30Hz) and Gamma band. (> 30 Hz) (NeuroSky, 2009). Valipour et al. has used FFT technique for the spectral analysis of alpha band during open and closed eyes (Valipour et al., 2013). Power Spectrum Densities (PSD) describes the power distribution of the signal as a function of frequency of a periodic signal. Saa et al. has analyzed the power spectral density during motor imagery for classifying the imaginary motor movements (Delgado, Saa & Gutierrez, 2010). (Rojas et al., 2016) has analyzed the EEG signal during imagined speech of Spanish vowels using digital signal processing techniques such as Filters, FFT and statistical analysis based on ANOVA test and has reported that by using signal processing it is possible to recognize the Spanish vowels from imagined speech. In general EEG signals are nonlinear and non-stationary. EEG data contains artifacts due to the surrounding noise. The amplitude of artifacts is quite large compared to the task related EEG signal. Artifacts due to eye blinks (electrooculography-EOG) and muscle movements (electromyography-EMG) has to be removed in order to analyze the signal for interpreting correct results. Independent Component Analysis (ICA) techniques have been used to remove the above-mentioned artifacts. There are many source separation algorithms designed for ICA. These techniques attempt to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

unmix the EEG signals into some number of underlying components that would result in clean EEG by nullifying the weight of unwanted components (Nolan et al., 2010; Jung et al., 2000).

Research in the domain of extracting significant information about brain connectivity from EEG recordings has been found to be active and still new findings are continuously reported. Cognitive researchers working in the field of neuroscience are actively involved in the estimation of brain connectivity of electrophysiological signals. In the study of brain connectivity, the central issue of determining directionality of the interactions among neural signals is of great interest because it allows revealing pathways of information flow within the nervous system. The brain connectivity can be defined in terms of anatomical connectivity (physical connections linking brain regions), functional connectivity (statistical dependence in spatially remote brain activity) and effective connectivity (the causal influence of one brain region on another i.e., directional relationships) (Friston, 2011; Greenblatt, 2012). The prevalence of brain connectivity estimators is useful in quantifying the frequency synchronization between the different regions of the brain. Various studies have been proposed in order to estimate functional synchronization and connectivity from EEG signals (Sakkalis, 2011). In recent research brain connectivity has been analyzed while imagining of speaking for similar and dissimilar sounds (Sandhya & Kavitha, 2015; Sandhya et al., 2015). Phase synchronization measures such as Phase Locking Value (PLV) is a functional connectivity parameter through which the phase coupling between two time series signals can be estimated using PLV (Lachaux et al., 1999). The whole brain functional connectivity has been investigated using Phase Locking Value for resting state signals (Schmidt et al., 2014).

In this work, EEG signals were acquired from healthy subjects while speaking and imagining of speaking Consonant-Vowel-Consonant (CVC) words using multiple trials in different time frames. Artifacts from the EEG signal were removed using ICA. Absolute powers were computed using FFT in each frequency band during the tasks for all the trials. Relative spectral power has been derived from absolute power values for each frequency band. This relative power analysis will be useful in contributing an effective EEG data analysis during speech imagery process. The EEG bands that will be dominant during speech and speech imagery can be recognized using this power spectral analysis which can be further extended to analyze the neuronal correlations during speech imagery tasks. The phase synchronization between two time series signals were estimated using PLV for multiple trials using CVC words. PLV will not only provide information about the dominating lobes of the brain during the speech imagery task but also about the connectivity among them.

METHODOLOGY

EEG signals were recorded in a sound proof room as per the guidelines of Institutional Ethics Committee of Department of Biomedical Engineering, SSN College of engineering, India. Two right-handed subjects participated in this work. All subjects were men and native Indians, between the mean age group of 19 and 22 (20.5 ± 2.12).

EEG Acquisition

EEG signals were acquired using EMOTIV EPOC, a 14 electrode wireless EEG acquisition device. Figure 1 shows EEG recording setup for a representative participant using EMOTIV EPOC. EEG signals were transferred to the system via Bluetooth. Signals from fronto-polar (AF3 and AF4), frontal (F3, F7,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. EEG recording set up

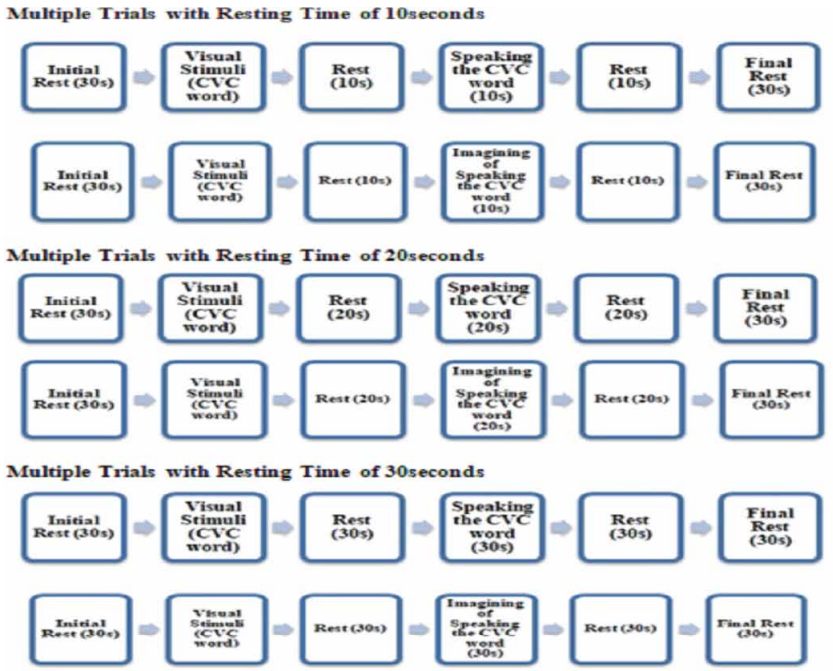
F4, F8), fronto-central (FC5 and FC6), temporal (T7, T8, P7, P8) and occipital (O1 and O2) lobes were acquired with a sampling rate of 128 Hz. These electrode locations are named according to standard 10-20 electrode placement system. The reference electrodes were chosen to be P3 and P4. EEG signals from EMOTIV EPOC were acquired within the in-built range of 0.001 to 43 Hz. Though the signals were acquired from all 14 electrodes, the region of interest related to speech and speech imagery are frontal and temporal lobes. So, the electrodes (F7, FC5, T7, P7) with respect to Broca's and Wernicke's regions located in left frontal and left temporal regions were analyzed exclusively, as these regions are responsible for speech production and speech comprehension.

Experimental Protocol

For each vowel in the vowel set 'a', 'e', 'i', 'o', 'u' one CVC word has been chosen. The chosen words are *'Can', 'Let', 'Did', 'Jog', 'Sun'*. These words were shown as visual stimuli to the subjects. The subjects were asked to speak the word and imagine of speaking the word silently in the mind without any articulation movement. An initial rest period of thirty seconds was provided to the subject on beginning of each trial and similarly a final rest period of 30seconds was given at the end of each trial. Different time frames were used while switching between speech and speech imagery process. The resting time for each protocol has been increased before and after the task with a time period of 10seconds, 20seconds and 30seconds respectively. Similarly, the task time with respect to speaking and imagining of speaking were also increased gradually with the time periods as mentioned above. Three sessions were carried out for each word. In the first session, subject was made to feel comfortable, and training had been given to understand the protocol. For the second session, subject was asked to speak the CVC word for five times. In the third session the subject was asked to imagine of speaking the CVC word for five times. This protocol of three different time frames was employed to test the difference in brain connectivity's as the brain was allowed to settle and switch over between speaking and imagining of speaking processes under varying time intervals. Five trials for each CVC word for every resting time of 10s, 20s and 30s.15 trials for one CVC word. For speech process the total number of trials is 75 (15*5 CVC word = 75) and similarly for speech imagery process the total of number of trials is 75 (15*5 CVC word = 75). The size of the EEG data for each CVC word in single trial with respect to 10s is (2 subjects* 14 channels * 900 samples), 20s (2subjects * 14 channels* 1800 samples) and 30s (2subjects* 14 channels * 3000 samples). Figure 2 shows the experimental protocol for speech and speech imagery process involving different time periods.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Experimental protocols for speech and speech imagery process involving different time periods



EEG Preprocessing

The acquired EEG signals were filtered between the frequency ranges of 0.5 Hz to 40 Hz using an IIR Band-Pass Chebychev filter for both speech and speech imagery signals (Subramanian et al., 2017). In order to remove the artifacts due to eye blinks and muscle movements the signals were further decomposed using Independent Component Analysis (ICA). ICA method is used for separating a multivariate signal into additive subcomponents. Each component is statistically independent from each other and non-Gaussian. A fast and stable non-linear ICA decomposition algorithm was implemented for automatic rejection of components related to artifacts (Delorme & Makeig 2004). After ICA decomposition, these signals were further filtered using smoothing filter. In this work moving average technique has been implemented for obtaining the smoothed signal. This method creates a series of averages of different subsets of the full data set and can be classified as a type of finite impulse response. Ultimately, the moving average function can be viewed as a low-pass filter (Azami et al., 2012). Using IIR band-pass chebyshev filter, the signals were segmented into EEG sub-band frequencies. The prominent EEG frequencies were found to be Delta (1-3 Hz), Theta (4-8Hz), Alpha (8 to 13 Hz), Beta (13 to 30 Hz) and Gamma (30 to 40 Hz). Power spectral density and relative powers were analyzed in each frequency band for EEG signals in multiple trials.

Power Spectral Density (PSD) and Relative Power

The frequency domain analysis was carried out using Fast Fourier Transform (FFT) in order to calculate the absolute power (μV^2 /Hz) and relative power (%). For a signal *x(t),* Fourier transform of the signal *x(t)* is given by *X(f).* The power spectral density is expressed as the square of the absolute values of FFT

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of the signal (i.e) power present in the signal as a function of frequency (i.e.) $|X(f)|^2=S_X(f)$. The absolute power of a band is the integral of all power values within the frequency range (Wang et al., 2016) (i.e.) The absolute spectral power is defined as the band of frequencies from f_0 Hz to f_1 Hz to the total power delivered at the output of a band-pass filter that passes all frequencies from f_0 Hz to f_1 Hz and stops other frequencies.

Absolute Spectral Power is given by:

$$\int_{-f_1}^{-f_0} S_X\left(f\right)df + \int_{f_0}^{f_1} S_X\left(f\right)df \tag{1}$$

Relative power measures the ratio of the total power in the band to the total power in the signal (i.e.) ratio of absolute power in each frequency band to the absolute power summed over the five frequency bands (Delta, Theta, Alpha, Beta, Gamma) (Yuvaraj et al., 2014).

Relative Spectral Power is given by:

$$\frac{\int_{-f_1}^{-f_0} S_X\left(f\right)df + \int_{f_0}^{f_1} S_X\left(f\right)df}{\int_{-\infty}^{\infty} S_X\left(f\right)df} \tag{2}$$

In this work relative power were analyzed in each frequency band for the multi trial speech and speech imagery EEG signals. Relative spectral powers were analyzed exclusively for the signals originated from the electrodes F7, FC5, T7 and P7. These electrodes were located near Broca's and Wernicke's area respectively.

Phase Locking Value (PLV)

Lachaux et al. (1999) demonstrated a practical method for the quantification of transient phase locking between two neuro electric signals. Phase Locking Value (PLV) is a phase synchronization measure used for bi-variate process. For each EEG electrode the phase time course is given by ϕ. The difference between two phase time course ($\Delta\phi$) quantifies the locking between the phases of given two electrodes. If the EEG signals of two electrodes rises and falls together during the given cognitive task then there is more synchronization (i.e) enhanced connectivity is observed between the given two electrodes and $\Delta\phi$ will be constant between trials. If de-synchronization or decreased connectivity is observed between two electrodes then $\Delta\phi$ will be random. PLV ranges between 0 and 1. PLV of 0 represents that two electrodes are not coupled whereas if PLV is 1 it indicates that one signal follows the other synchronously and the two electrodes have a strong connectivity. Phase Locking Value (PLV) for a single trial is given by:

$$PLV\left(t,f\right) = \frac{1}{N}\left|\sum_{n=1}^{N}\exp\left(j\left\{\Delta\Phi_n\left(t,f\right)\right\}\right)\right| \tag{3}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

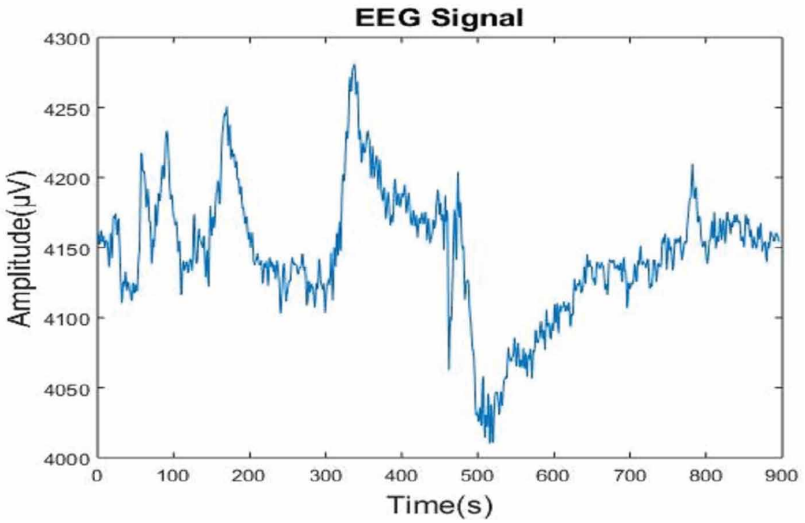
where N is the number of samples, $\Delta\Phi_n(t,f) = \Phi_1(t,f) - \Phi_2(t,f)$ is the instantaneous phase difference. In case of multiple trials N is the number of trials and PLV was computed over multiple trials at every instance of time. In this work, PLV has been computed between various electrodes for each EEG frequency sub-bands while speaking and imagining of speaking the given CVC words. Importance was given to inter-hemisphere and intra-hemisphere frontal and temporal lobe electrodes due to their close proximity to the Broca's area and Wernicke's area.

RESULTS AND DISCUSSION

Relative Power Analysis While Speaking the CVC Word

Frontal and fronto-central electrodes (F3, F7, FC5, F4, F8, FC6) with respect to left and right hemisphere of the brain were analyzed individually during speaking of the CVC word *'can'*. F7 and FC5 electrodes are located near Broca's area of left frontal lobe which is responsible for speech production. The signal analysis for F7 electrode has been discussed in detail for understanding the nature of the signal while speaking the representative CVC word. Similarly, other frontal electrodes were also analyzed with the same procedure. Figure 3 shows the F7 electrode EEG signal while speaking the CVC word *'can'* with a resting time of 10 seconds. It can be observed from Figure 3 that amplitude ranges are very high. The signals recorded using Emotiv Epoc device was transmitted via Bluetooth and hence the amplitude ranges are very high. DC level of the signal occurs at approximately 4200 μV, negative voltages are transmitted as positive values less than the average level, and positive voltages are transmitted as positive values greater than the average. Before performing any kind of analysis such as FFT, it is necessary to apply some kind of DC offset removal.. The simplest method is to normalize the signal by subtracting the average from the entire data channel. The normalized signal is further filtered using bandpass chebyshev filter in a frequency range of 0.5 Hz to 40 Hz.

Figure 3. EEG signal-F7 electrode



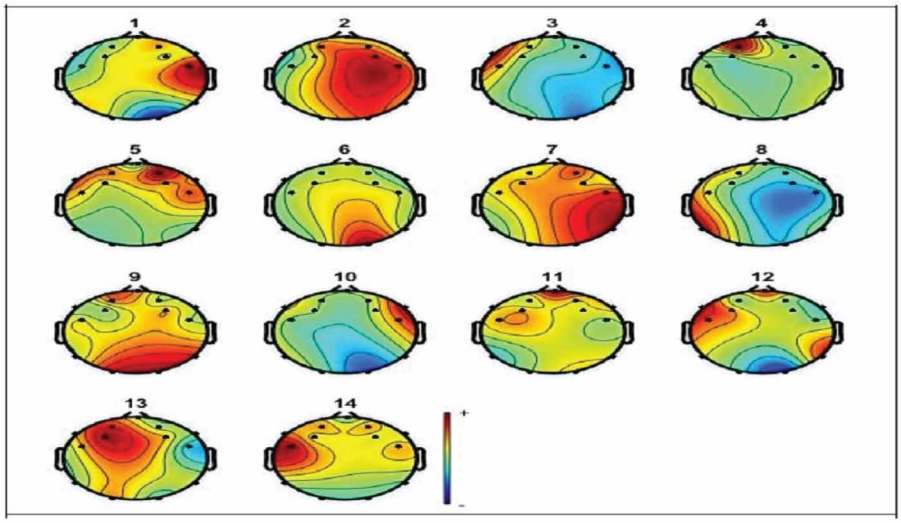
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

EEG signals are often affected by artifacts due to the surrounding sources such as background noise. Removal of artifacts is a major step in preprocessing stage. In order to remove these artifacts, the filtered EEG signals were decomposed using ICA. The output of ICA decomposition is a data activation matrix. The channel data are arranged in a matrix of *n* channels (rows) by *t* time points (columns). Each row in the matrix represents the time course activity of one component that is spatially filtered from the channel data. The scale of each component activity time course uses arbitrary units of the original data. Components with eye and muscle artifacts were rejected based on the power spectrum. Components with smoothly decreasing EEG spectrum were considered to be a typical artifact. ICA obtained from EEG signal while speaking the CVC word *'can'* is shown in Figure 4. Red colour indicates maximum synchronous activity (maximum range of the scale) and blue colour indicates non-synchronous activity (negative range of scale) and the less active components are represented in mid-range of the scale (i.e.) mix of yellow and green colour. From Figure 4 it has been observed that ICA components with respect to IC2, IC4, IC7, IC11, IC12 and IC13 were showing synchronous activity during speaking of the CVC word *'can'*. Similarly, ICA components with respect to IC1, IC3, IC5, IC6, IC8 and IC10 were showing a non-synchronous activity. ICA components IC9 and IC14 were showing a mild activity. After the rejection of ICA components related to artifacts, a total of 12 components were selected for further analysis.

After ICA decomposition, the signal has been further filtered using smoothing filter. The main function of the smoothing filter is to capture the important patterns in data and also used for filtering out unnecessary noise. The smoothed signal was obtained using moving average technique. Signals were further segmented into EEG frequency bands such as delta, theta, alpha, beta and gamma bands using band-pass chebyshev filter and their corresponding power spectral densities were estimated in each frequency band using FFT. The prominent EEG frequencies are Delta (1-3 Hz), Theta (4-8Hz), Alpha (8 to 13 Hz), Beta (13 to 30 Hz) and Gamma (30 to 40 Hz).

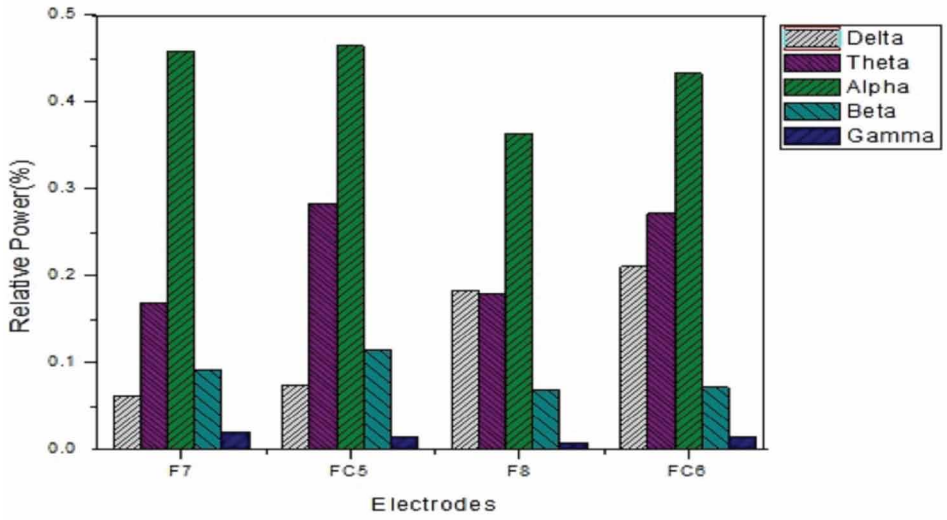
Based on the absolute power for each band, the relative power has been computed for electrodes with respect to frontal electrodes in left and right hemisphere of the brain. Figure 5 shows the relative power for frontal electrodes (F7, FC5, F8, and FC6) while speaking the CVC word *'can'* in a time period of 10

Figure 4. ICA components obtained from EEG signal while speaking the CVC word 'can'



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Relative power of frontal electrodes during speech process



seconds. From Figure 5 it was observed that relative power of alpha bands were found to be high for F7 and FC5 electrode. It was also observed that relative power of delta, theta, beta and gamma bands were low compared to alpha bands during speech task. It was also observed that right hemisphere electrodes (F8, FC6) were found to have a lower relative power compared to the left hemisphere electrodes (F7, FC5).

Relative Power Analysis While Imagining of Speaking the CVC Word

Temporal electrodes (T7, T8, P7, P8) with respect to left and right hemisphere of the brain were analyzed individually during imagining of speaking the CVC word *'can.'* T7 electrode was chosen for speech imagery analysis as this electrode is located near the Wernicke's area of temporal lobe. Wernicke's area is responsible for thinking and speech perception. The signal analysis for T7 electrode has been discussed in detail for understanding the nature of the signal while imagining of speaking the representative word. Similarly, for other temporal electrodes were also analyzed with the same procedure. As discussed for the analysis of speech EEG signals in previous sections, the same procedure will be followed for speech imagery signals. The filtered EEG signals were decomposed using ICA in order to remove the artifacts. The smoothed signals were further segmented into EEG frequency bands such as delta, theta, alpha, beta and gamma bands using band-pass chebyshev filter and their corresponding power spectral densities were estimated in each frequency band using FFT. Based on the absolute power for each band, the relative power has been computed for electrodes with respect to temporal electrodes in left and right hemisphere of the brain. Figure 6 shows the relative power for temporal electrodes (T7, T8, P7, P8) while imagining of speaking the CVC word *'can'* in a time period of 10 seconds. T7 and P7 electrode is located near Wernicke's area of left temporal lobe. From Figure 6 it was observed that relative power of theta band were found to be higher for T7 and P7 electrode while imagining of speaking the CVC word *'can'*. It was also observed that relative power of delta, alpha, beta and gamma bands were low compared to theta band during speech imagery task. It was also observed that right hemisphere electrodes (T8, P8) were found to have a lower relative power values compared to the left hemisphere electrodes (T7, P7).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Relative power for temporal electrodes during speech imagery process

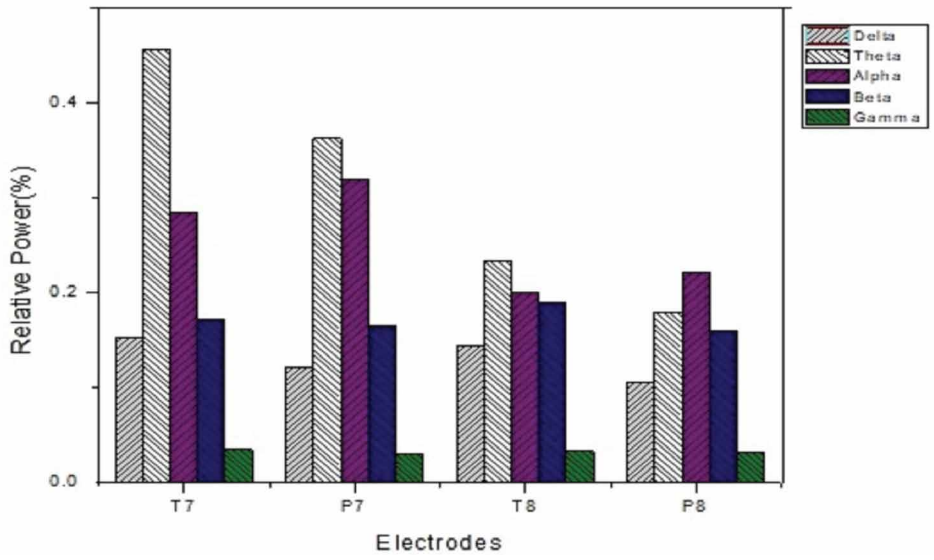


Figure 7 shows the averaged relative power of F7 electrode in 10 seconds, 20 seconds and 30 seconds of alpha band while speaking the CVC word *'can'*. Figure 8 shows the averaged relative power of T7 electrode in 10 seconds, 20 seconds and 30 seconds of theta band while imagining of speaking the CVC word *'can'*. In each trial it was evident that alpha and theta were found to have a high relative power value while speaking and imagining of speaking a CVC word. From Figure 7 and Figure 8 it has also been observed that, relative power value was progressively increasing with the increasing resting time period for speech and speech imagery protocols. It was also observed that right hemisphere electrodes (F4, F8, FC6, T8,P8) were found to have a lower relative power values compared to the left hemisphere electrodes (F3, F7, FC5, T7, P7).

PLV Analysis for Speech Imagery Process Involving CVC Words in Multiple Trials

PLV has been estimated for frontal (F7, F3, F4, F8), fronto-central (FC5, FC6), temporal electrodes (T7, T8, P7, P8) while speaking the CVC word in multiple trials. A 10*10 matrix were obtained for the frontal, fronto-central and temporal electrode combinations. These matrixes were obtained for all five sub-band EEG frequencies (delta, theta, alpha, beta and gamma). PLV has been estimated during speaking of a CVC word with respect to 10s resting time, 20s resting time and 30s resting time individually for each EEG bands. In all cases it has been found that alpha band was dominant band compared to other bands. Figure 9 shows the averaged PLV connectivity during speaking of the CVC word *'Can'* in alpha band with resting time of 10s with respect to lateral, medial, coronal and axial view of the brain in multiple trials. The electrodes are represented in terms of Brodmann areas. Frontal electrodes are represented as F7-47L, F8-47R, F3-8L, F4-8R (47 L&R-Inferior Prefrontal Gyrus, 8 L&R-Frontal Eye Field with respect to left and right hemisphere of the brain). Similarly fronto-central electrodes are represented as FC5-45L, FC6-45R (45 L &R-Broca's area of left and right hemisphere) followed temporal electrodes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Average relative power for speech process in alpha band under three different resting time periods for F7 electrode

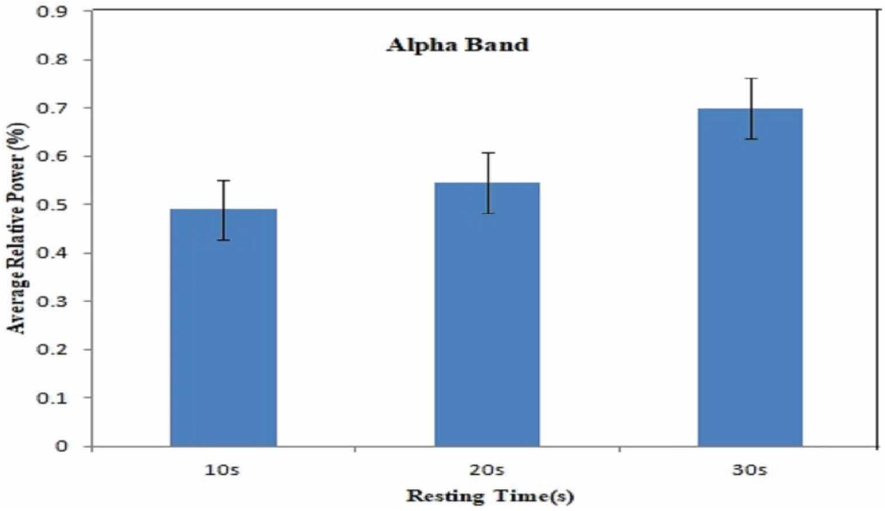
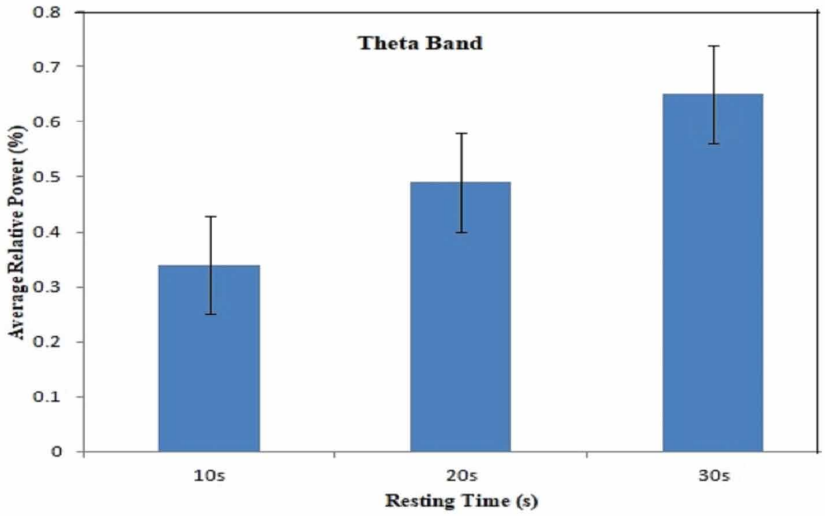


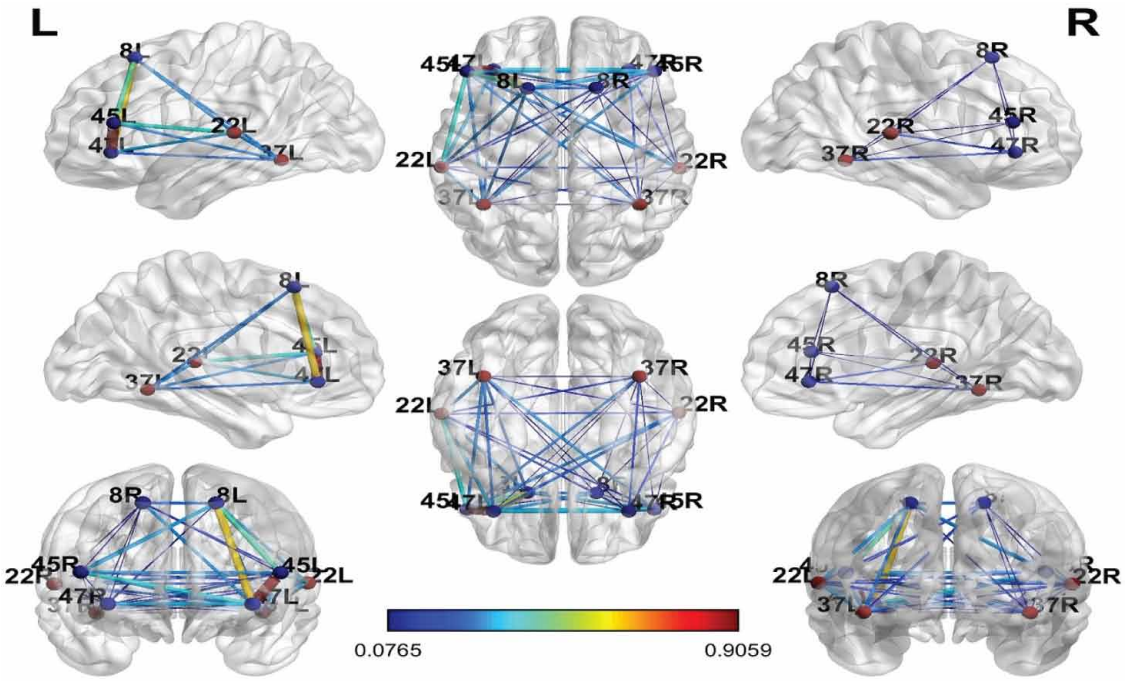
Figure 8. Average relative power for speech imagery process in theta band under three different resting time periods forT7 electrode



as T7-22L, T8-22R, P7-37L, P8-37R (22 L&R- Wernicke's area of left and right hemisphere, 37 L&R-Fusiform Gyrus of left and right hemisphere of the brain). From Figure 9 it was observed that the left hemisphere electrodes with respect to frontal and fronto-central electrode combination (F7 with FC5) showed maximum phase lock value which indicates that the difference between these two electrodes quantifies the locking (i.e) both the signals phases are locked and one signal follows the other perfectly (rise and fall time of both the electrodes were in phase). Scale ranges from blue to maroon color. Blue indicates the PLV value of 0.1 and below. Maroon color indicates higher PLV connectivity between the electrodes of value 0.9 and above.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Average PLV connectivity during speaking of the CVC word with resting time of 10s in multiple trials



The connectivity matrix of averaged PLV in multiple trials is shown in Figure 10. From Figure 10, the connectivity between F7 with FC5 was found to be 0.9 which was the highest compared to other electrode combinations. The inter hemispheric electrode combinations showed a lower connectivity during speaking of the CVC word *'Can'*. Scale ranges from blue to yellow color. Blue indicates lower connectivity and yellow indicates higher connectivity between the electrodes. Lower connectivity can be observed between F7 with P8 and F3 with P8 electrode combinations. The connectivity between pairs of temporal electrodes were low compared to that of frontal and frontal-central electrode combinations which proves that Broca's area were active during speech production.

Figure 11 shows the averaged PLV connectivity during speaking of the CVC word *'Can'* in alpha band with resting time of 20s with respect to lateral, medial, coronal and axial view of the brain in multiple trials. From Figure 11 it was evident that the left hemisphere electrodes with respect to frontal and fronto-central electrode combination (F7 with FC5) showed a high phase lock value compared to other electrode combinations. PLV was found to be increasing in frontal and fronto central electrodes when the resting time of the task is increased from 10s to 20s.

Figure 12 shows the averaged PLV connectivity during speaking of the CVC word *'Can'* in alpha band with resting time of 30s with respect to lateral, medial, coronal and axial view of the brain in multiple trials. From Figure 12 it was evident that the left hemisphere electrodes with respect to frontal and fronto-central electrode combination (F7 with FC5) showed a maximum phase lock value compared to other electrode combinations. PLV was found to be increasing in frontal and fronto central electrodes when the resting time of the task is increased from 20s to 30s which is evident that the electrode F7 and FC5 signals are more active while speaking a CVC word when the resting time of the task is increased.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Average PLV connectivity matrix with a resting time of 10s while speaking a CVC word in multiple trials

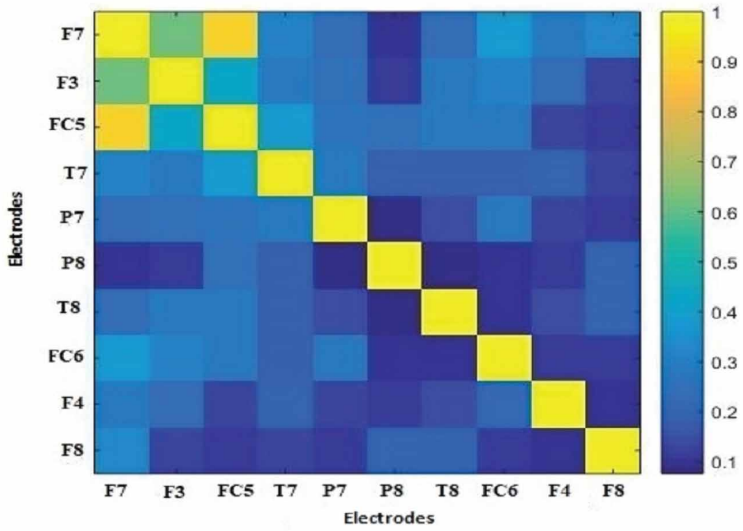
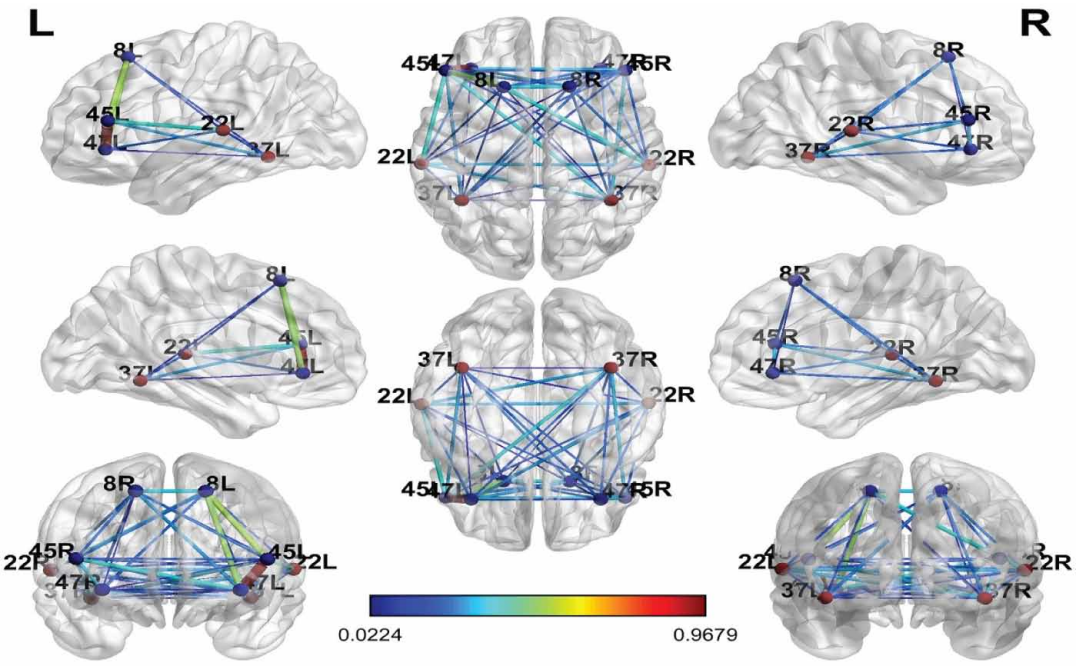
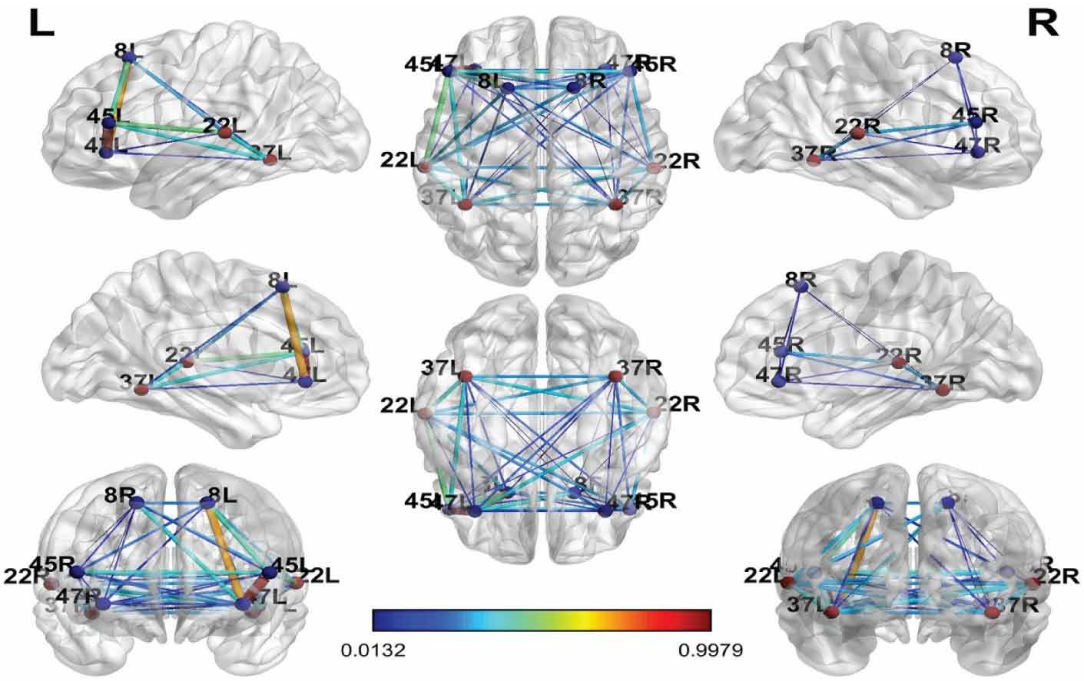


Figure 11. Average PLV connectivity during speaking of a CVC word with resting time of 20s in multiple trials



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Average PLV connectivity during speaking of the CVC word with resting time of 30s in multiple trials



PLV Analysis for Speech Imagery Process Involving CVC Words -Single and Multiple Trials

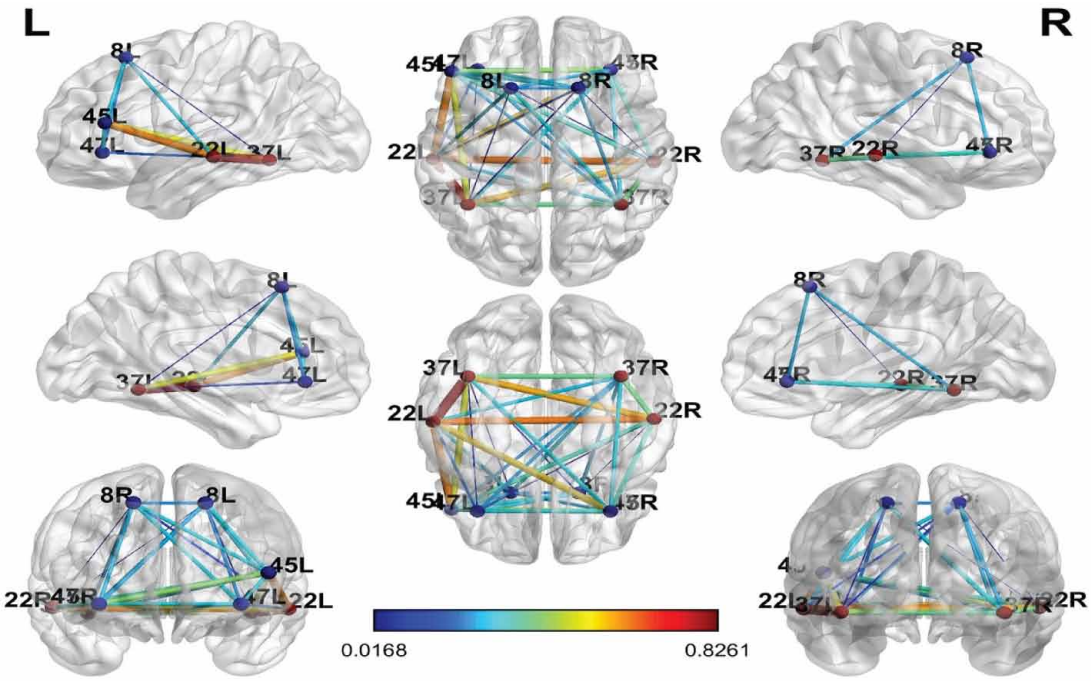
PLV has been estimated for frontal (F7, F3, F4, F8), fronto-central (FC5, FC6), temporal electrodes (T7, T8, P7, P8) while imagining of speaking a CVC word in multiple trials. A 10*10 matrix were obtained for the frontal, fronto-central and temporal electrode combinations. PLV has been computed while imagining of speaking a CVC word with respect to 10s resting time, 20s resting time and 30s resting time individually in multiple trials for each EEG bands.

In all cases it has been found that theta band was the dominant band compared to other bands. Figure 13 shows the averaged PLV connectivity during imagining of speaking of the CVC word *'Can'* in theta band with resting time of 10s with respect to lateral, medial, coronal and axial view of the brain in multiple trials. The electrodes are represented in terms of Brodmann'areas. Frontal electrodes are represented as F7-47L, F8-47R, F3-8L, F4-8R (47 L&R-Inferior Prefrontal Gyrus, 8 L&R-Frontal Eye Field with respect to left and right hemisphere of the brain). Similarly, fronto-central electrodes are represented as FC5-45L, FC6-45R (45 L &R-Broca's area of left and right hemisphere) followed temporal electrodes as T7-22L, T8-22R, P7-37L, P8-37R (22 L&R- Wernicke's area of left and right hemisphere, 37 L&R-Fusiform Gyrus of left and right hemisphere of the brain).

From Figure 13 it was observed that the left hemisphere electrodes with respect to temporal electrode combination (T7 with P7) showed maximum phase lock value which indicates that the difference between these two electrodes quantifies the locking (i.e.) both the signals phases are locked and one signal follows the other perfectly (rise and fall time of both the electrodes were in phase). Scale ranges from blue

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

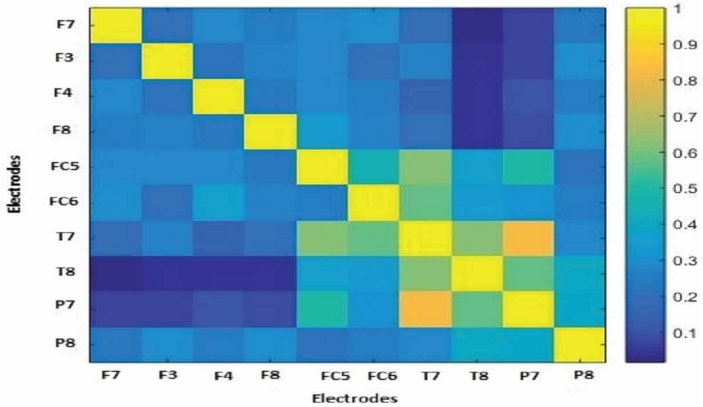
Figure 13. Average PLV connectivity during imagining of speaking the CVC word with resting time of 10s in multiple trials



to maroon color. Blue indicates the PLV value of 0.1 and below. Maroon color indicates higher PLV connectivity between the electrodes of value 0.82 and above.

The connectivity matrix of averaged PLV in multiple trials is shown in Figure 14. From Figure 14, it has been observed that the connectivity between T7 and P7 was found to be 0.82 which was the highest value compared to other electrode combinations. The inter hemispheric electrode combinations showed a

Figure 14. Average PLV connectivity matrix with a resting time of 10s while imagining of speaking the CVC word in multiple trials



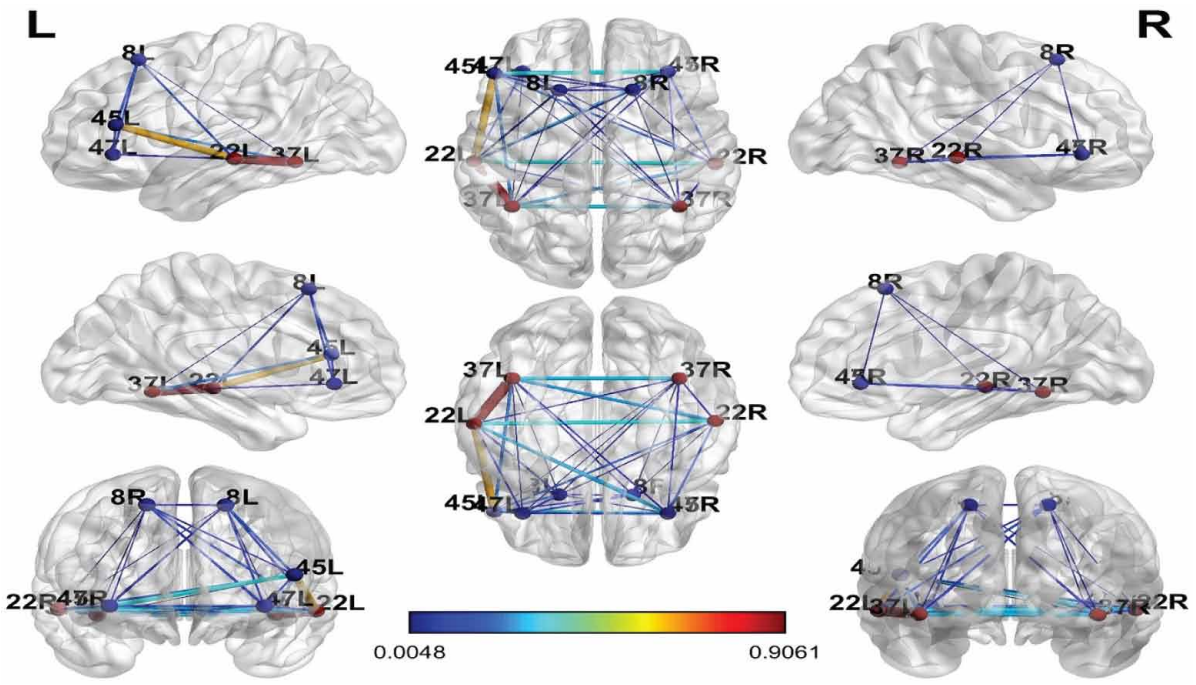
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

lower connectivity during imagining of speaking the CVC word *'Can'*. Scale ranges from blue to yellow color. Blue indicates lower connectivity and yellow indicates higher connectivity between the electrodes. Lower connectivity can be observed between frontal electrode combinations. The connectivity between temporal electrodes was high compared to that of frontal and frontal-central electrode combinations which prove that Wernicke's area was active during speech imagery process.

Figure 15 shows the averaged PLV connectivity during imagining of speaking of the CVC word *'Can'* in theta band with resting time of 20s with respect to lateral, medial, coronal and axial view of the brain in multiple trials. From Figure 15 it was evident that the left hemisphere electrodes with respect to temporal electrode combination (T7 with P7) showed a high phase lock value compared to other electrode combinations. PLV was found to be increasing in temporal electrodes when the resting time of the task is increased from 10s to 20s.

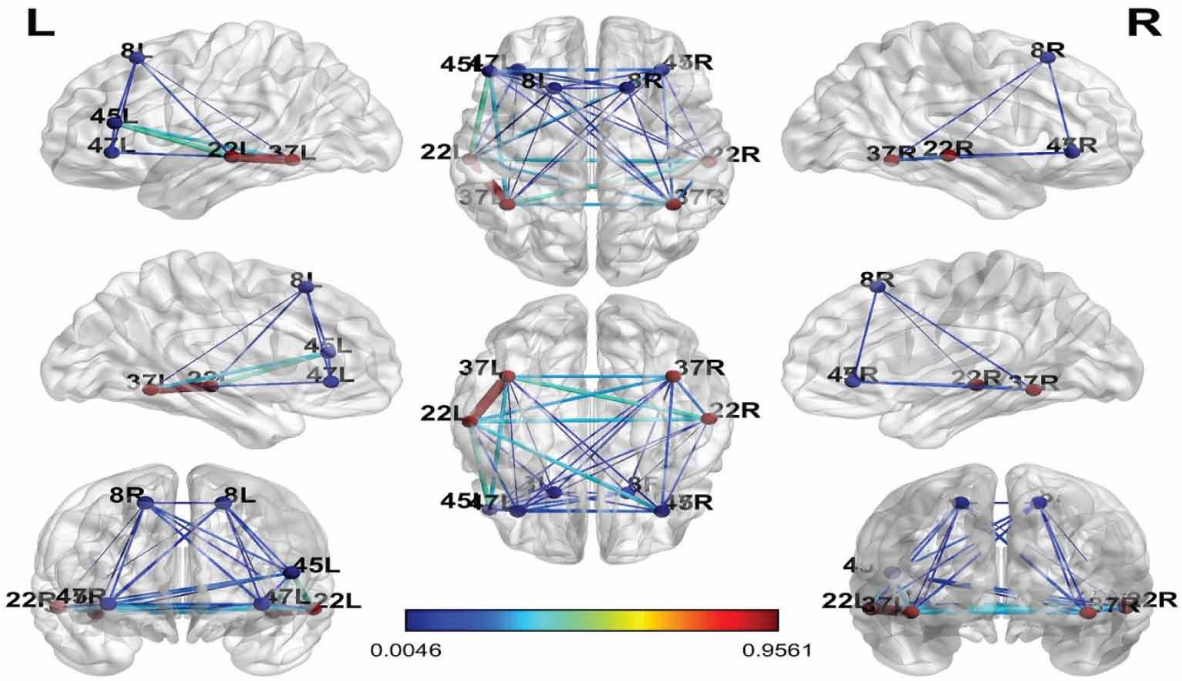
Figure 16 shows the PLV connectivity during imagining of speaking of the CVC word *'Can'* in theta band with resting time of 30s with respect to lateral, medial, coronal and axial view of the brain in multiple trials. From Figure 16 it was evident that the left hemisphere electrodes with respect to temporal combination (T7 with P7) showed a maximum phase lock value compared to other electrode combinations. PLV was found to be increasing in temporal electrodes when the resting time of the task is increased from 20s to 30s which is evident that the electrodes T7 and P7 signals are more active while imagining of speaking a CVC word when the resting time of the task is increased.

Figure 15. Average PLV connectivity during imagining of speaking of the CVC word with resting time of 20s in multiple trials



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. Average PLV connectivity during imagining of speaking the CVC word with resting time of 30s in multiple trials



CONCLUSION

Fast Fourier transform has been used in this paper to analyze signals in the frequency domain for speech and speech imagery. It extracts the information present in the frequencies of a particular band and the amount of computation is reduced compared to other transforms. Power spectrum densities of each frequency band were derived based on the absolute values of FFT. Relative power for each inter and intra hemispheric frontal and temporal electrodes in each frequency band were analyzed exclusively. It was observed that relative power of theta band was dominant during speech imagery process and relative power of alpha was active during speech process. It was also observed that the relative power value was found to be increasing with the increasing resting time period involving different protocols. Functional brain connectivity parameters such as Phase Locking Value have been estimated for speech and speech imagery processes involving CVC words in multiple trials of different time periods in units of ten. PLV reveal the degree of synchronicity between two regions of the brain when a task is performed. It was observed that PLV of left hemispheric frontal electrode combination (F7 with FC5) (> 0.85) showed maximum phase lock value in alpha band during speech process which indicates that the difference between these two electrodes quantifies the locking (i.e.) both the signals phases are locked and one signal follows the other perfectly. Similarly, PLV has been computed for speech imagery process using CVC words in multiple trials. It was inferred that left hemispheric temporal electrode combination (T7 with P7) showed maximum phase lock value in theta band (>0.9) during speech imagery process. PLV values were found to increase progressively with the increasing resting time in units of ten. Results confirmed that left frontal electrodes were active during speaking of a CVC word and left temporal electrodes were

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

active during imagining of speaking a CVC word. This work can be extended by including Consonant-Vowel-Consonant –Consonant (CVCC) words for each vowel, and also by including different brain connectivity parameters for analysis. Therefore, through these adaptive signal processing methods and brain connectivity estimators, future scope of this work can be used to develop a control scheme for speech prostheses for the production of words in speech impaired people using speech imagery process.

REFERENCES

Azami, H., Mohammadi, K., & Bozorgtabar, B. (2012). An improved signal segmentation using moving average and Savitzky-Golay filter. *Journal of Signal and Information Processing*, *3*(01), 39–44. doi:10.4236/jsip.2012.31006

Callan, D. E., Callan, A. M., Honda, K., & Masaki, S. (2000). Single-sweep EEG analysis of neural processes underlying perception and production of vowels. *Brain Research. Cognitive Brain Research*, *10*(1-2), 173–176. doi:10.1016/S0926-6410(00)00025-2 PMID:10978705

Chengaiyan, S., & Anandhan, K. (2015). Analysis of speech imagery using functional and effective EEG based brain connectivity parameters. *International Journal of Cognitive Informatics and Natural Intelligence*, *9*(4), 33–48. doi:10.4018/IJCINI.2015100103

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi:10.1016/j.jneumeth.2003.10.009 PMID:15102499

Dogil, G., Ackermann, H., Grodd, W., Haider, H., Kamp, H., Mayer, J., & Wildgruber, D. (2002). The speaking brain: A tutorial introduction to fMRI experiments in the production of speech, prosody and syntax. *Journal of Neurolinguistics*, *15*(1), 59–90. doi:10.1016/S0911-6044(00)00021-X

Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, *1*(1), 13–36. doi:10.1089/brain.2011.0008 PMID:22432952

Greenblatt, R. E., Pflieger, M. E., & Ossadtchi, A. E. (2012). Connectivity measures applied to human brain electrophysiological data. *Journal of Neuroscience Methods*, *207*(1), 1–16. doi:10.1016/j.jneumeth.2012.02.025 PMID:22426415

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*(5), 393–402. doi:10.1038/nrn2113 PMID:17431404

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163–178. doi:10.1111/1469-8986.3720163 PMID:10731767

Kosslyn, S. M., & Pylyshyn, Z. (1994). Image and brain: The resolution of the imagery debate. *Nature*, *372*(6503), 289–289.

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, *8*(4), 194–208. doi:10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C PMID:10619414

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36. doi:10.1016/0010-0277(85)90021-6 PMID:4075760

NeuroSky. (2009). Brain wave signal (EEG) of NeuroSky, Inc. Publisher: NeuroSky Brain-Computer Interface Technologies. Retrieved from http://www.frontiernerds.com/files/neurosky-vs-medical-eeg.pdf

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162. doi:10.1016/j.jneumeth.2010.07.015 PMID:20654646

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Loevenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*, *261*, 220–239. doi:10.1016/j.bbr.2013.12.034 PMID:24412278

Price, C. J., Crinion, J., & MacSweeney, M. (2011). A generative model of speech production in Broca's and Wernicke's areas. *Frontiers in Psychology*, *2*, 237. doi:10.3389/fpsyg.2011.00237 PMID:21954392

Rojas, D. A., Góngora, L. A., & Ramos, O. L. (2016). EEG signal analysis related to speech process through BCI device EMOTIV, FFT and statistical methods. *Journal of Engineering and Applied Sciences (Asian Research Publishing Network)*, 3074–3080.

Saa, J. F. D., & Gutierrez, M. S. (2010, June). EEG signal classification using power spectral features and linear discriminant analysis: a brain computer interface application. In *Eighth Latin American and Caribbean Conference for Engineering and Technology* (pp. 1-7). Arequipa: LACCEI.

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Computers in Biology and Medicine*, *41*(12), 1110–1117. doi:10.1016/j.compbiomed.2011.06.020 PMID:21794851

Sandhya, C., Srinidhi, G., Vaishali, R., Visali, M., & Kavitha, A. (2015, July). Analysis of speech imagery using brain connectivity estimators. In *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)* (pp. 352-359). IEEE. 10.1109/ICCI-CC.2015.7259410

Schmidt, B. T., Ghuman, A. S., & Huppert, T. J. (2014). Whole brain functional connectivity using phase locking measures of resting state magnetoencephalography. *Frontiers in Neuroscience*, *8*, 141. doi:10.3389/fnins.2014.00141 PMID:25018690

Subramanian, M., Geethanjali, B., Seshadri, N. G., Venkat, B., & Vijayalakshmi, R. (2017). Visualization of Brain Activation During Attention-Demanding Tasks Using Cognitive Signal Processing. *International Journal of Cognitive Informatics and Natural Intelligence*, *11*(1), 60–81. doi:10.4018/IJCINI.2017010105

Teplan, M. (2002). Fundamental of EEG Measurement. *Measurement Science Review*, *2*, 1–11.

Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, *6*, 314. doi:10.3389/fnhum.2012.00314 PMID:23226121

Valipour, S., Shaligram, A. D., & Kulkarni, G. R. (2013). Spectral analysis of EEG signal for detection of alpha rhythm with open and closed eyes. *Int. J. Eng. Innovative Technol*, *3*(6), 1–4.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Wang, Y., Sokhadze, E. M., El-Baz, A. S., Li, X., Sears, L., Casanova, M. F., & Tasman, A. (2016). Relative power of specific EEG bands and their ratios during neurofeedback training in children with autism spectrum disorder. *Frontiers in Human Neuroscience*, *9*, 723. doi:10.3389/fnhum.2015.00723 PMID:26834615

Yuvaraj, R., Murugappan, M., Ibrahim, N. M., Omar, M. I., Sundaraj, K., Mohamad, K., & Satiyan, M. (2014). On the analysis of EEG power, frequency and asymmetry in Parkinson's disease during emotion processing. *Behavioral and Brain Functions*, *10*(1), 12. doi:10.1186/1744-9081-10-12 PMID:24716619

This research was previously published in the International Journal of Software Science and Computational Intelligence (IJSSCI), 10(4); edited by Mehdi Khosrow-Pour, D.B.A.; pages 44-61; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 83
Nursing Students' Perception of Medical Information Protection in Hospitals

Hyeon-Cheol Jeong
Sahmyook University, Korea

ABSTRACT

The purpose of this article is to investigate the perception of medical information protection. The subjects were 281 students from 2 nursing students from September 1, 2015 to October 20, 2015. A total of 28 self - administered questionnaires were used. As a result, the perception of medical information protection behavior was 3.39 ± 0.51. In the fourth grade, the 'privacy protection' category was low (p = .007) and the 'communication' category was high (p = .045). In case of having experience in practice, 'privacy protection' category (p = .010) and 'medical information management' category (p = .041) were low and 'communication' (p = .021) was high. In the 'communication' category, the perception score was high when the medical information protection law (p = .001) and the nurses' code of ethics (p = .040) were known. When the subject completed 'medical security education' (p = .004), the perception was high. In conclusion, it is necessary for schools and medical institutions to develop nursing education programs for their educational environment and role.

1. INTRODUCTION

In the early 1990s, the order communication system (OCS) was introduced; later on, in the late 1990s, the picture archiving communication System (PACS), which is based on the hospital information system (HIS), was introduced on a large scale. Medical informatization became more systematic in the 2000s, with many hospitals introducing electronic medical record (EMR) systems that make it possible to input, save, and exchange patient information (Suh & Lee, 2014).

EMR is the input, organizing, and archiving of the patient's medical care, surgery, and examination records based on computerization in the form of promised codes. As a result of examining the EMR

DOI: 10.4018/978-1-7998-1204-3.ch083

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

system for medical institutions in Korea in 2010, the rate of introduction of EMR reached 37.2%, and the rate of introduction of tertiary medical institutions was 50.2%, which is higher than the rate of introduction of US medical institutions (Jin & Choi, 2012).

In the medical environment, informatization converts the existing analog medical information into digital and creates a medical information network to improve the quality and efficiency of medical service (Shick, 1998). Such medical information includes personal information (eg, location, family, assets, social life, sex life, personality and daily habits) and very sensitive content (eg, injuries, illnesses, and treatment records) as well as the patient's health status.

Therefore, if medical information is exposed to a third party, it might violate the fundamental rights of privacy and freedom stipulated in the Constitution. Additionally, it might violate human's most fundamental rights, such as that of human dignity and values, and the right to pursue happiness. In particular, EMRs can easily be copied and sent to others. For this reason, if they are leaked and any modified and fabricated information is reproduced, the concerned individuals' privacy is likely to be violated much more than when dealing with conventional paper-based information. Therefore, it seems important to treat EMRs very carefully (Kim, 1998; Jeun, 2013; Koskosas & Asimopoulos, 2011).

Nurses, who care for patients 24 hours per day at hospitals and can easily access patients' personal information, are often in the most critical position for protecting patients' medical information. Nursing university students, as future nurses, are also highly involved in protecting patients' medical information during their period of clinical practice (Jung & Jung, 2011). Nevertheless, what nursing university students are taught about protection of patients' personal information differs widely depending on their school; there is not, at present, any standardization of such education. During their clinical practice courses, these students are given an ID allowing them to access patients' personal information, and thus the protection of such information is dependent on each student's personal ethics (Kim, Jeong & Song, 2013).

Despite the rapid informatization of the medical environment, there does not appear to be much education or research on nursing university students' own actions to protect patients' medical information. Therefore, this study attempts to identify students' perception of how to protect patients' medical information, and thereby provides fundamental material for developing an education program for students to protect patients' information in this rapidly changing medical environment.

2. STUDY METHOD

2.1. Study Design

This is a descriptive survey study on nursing university students' perceptions of the protection and management of patient information.

2.2. Study Subjects

The study subjects were students from the nursing departments of two universities in Seoul. The minimum sample size was determined to be 210 subjects using the G*power 3.1.5 program, according to an α of .05, a power of .95, and an effect size (F) of 0.5. A total of 350 questionnaires were distributed, with consideration of non-participation, and 294 copies (84%) were collected. Of these collected copies, 13 with insincere answers were excluded. Finally, 281 copies were used for data analysis.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2.3. Study Tools

The questionnaire used in this study was developed with reference to the questionnaire developed by Jung & Jung (2011), which was used to study nursing university students' protective actions concerning patients' medical information, as well as the 4th revision of the Korean Code of Nursing Ethics and the 1st revision of the Korean Ethical Guidelines of Nursing of the Korean Nurses Association. In particular, drawing on Article 10 (Privacy Protection), Article 11 (Confidentiality), and Article 12 (Management Responsibilities of Medical Records) in Chapter 3 (Ethics for Patients) of the Korean Code of Nursing Ethics, we modified and supplemented the styles and content adequacy.

Given that the subjects of this study were nursing students, we conducted the pilot test with five students at the nursing department and two new nurses. Based on the pilot subjects' ratings of the adequacy of the phrasing, understanding of the content, and time to complete the questionnaire, we selected 28 questions for further testing. The content validity of these questions was examined by three professors from the nursing department and two head nurses with a Ph.D. degree who were working at the university hospital. They determined that all 28 questions were valid (CVI = 0.9 or more). As a result, all questions were incorporated into the questionnaire without any changes.

This questionnaire comprises 28 questions in four categories to measure the degree of perception of patient information protection ("privacy protection," 4 questions; "confidentiality," 8 questions; "medical information management," 9 questions; and "communication," 7 questions). In this questionnaire, each question is rated on a 5-point Likert scale. Higher scores indicate a more positive perception of the protective actions of a patient's medical information.

When Jeong developed the original measurement tool, its reliability (Cronbach's α) was .90 (Jung & Jung, 2011). In this study, the Cronbach's α was .91.

2.4. Data Collection Period and Method

We collected data from Sep. 1 to Oct. 20, 2015. Before data collection, the researcher was given permission from the dean of each nursing department and explained the purpose of the study to subjects. All data were collected when subjects did not have a class. The students who determined to participate were given a self-administered questionnaire, and each questionnaire copy was collected by a student representative. The questionnaire contained an introduction to the researcher and described the purpose of this study, as well as a written explanation that the questionnaire would not be used for anything other than the stated study purpose, that participants could withdraw their participation at any time.

2.5. Data Analysis

For data analysis, SPSS Statistics (Ver. 21) was used:

1. Subjects' perception of the actions for protecting patients' medical information was analyzed using descriptive statistics, including minimum and maximum values, means, and standard deviations;
2. Differences in subjects' perceptions according to their general characteristics were analyzed using t-tests;
3. The correlations between the categories of perception of actions to protect patients' medical information were analyzed using Pearson's correlation coefficients.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3. STUDY RESULTS

3.1. Characteristics of Study Subjects

The general characteristics of the study subjects are presented in Table 1. Their average age was 21.4 years, and most were senior students (at 41.6%).

The majority (59.4%) had experienced hospital practice and were aware (66.5%) of the Medical Information Protection Act. Furthermore, 84.0% were aware of the Korean Code of Nursing Ethics developed by the Korean Nurses Association. However, 32.4% did not have a medical information security education. Regarding education on the protection of patients' personal information, the largest proportion (29.1%) had received such education in their fundamental nursing education class.

3.2. Perception About the Protection of Medical Information

Table 2 shows the results of analyzing the perception level of medical information protection for patients. The total score for this perception questionnaire was 3.39 ± 0.51 (5-point scale). With regard to category, confidentiality (4.20 ± 0.58) and communication (4.20 ± 0.65) had the highest scores.

3.3. Scores of Questions on Protective Actions for Patients' Medical Information

Table 3 shows the scores for each question of the questionnaire on perceptions of protective actions for patients' medical information, grouped by category.

Table 1. General characteristics of subjects (N = 281)

Characteristics	Classification	n or Mean	% or ±SD
Age		21.4	±2.17
Grade	1	51	18.1
	2	64	22.8
	3	49	17.4
	4	117	41.6
Hospital practice experience	Yes	167	59.4
	No	114	40.6
Are you aware of the Medical Information Protection Act?	Yes	187	66.5
	No	94	33.5
Are you aware of the Korean Code of Nursing Ethics?	Yes	236	84.0
	No	45	16.0
Received education on medical information security	No	129	32.4
	In fundamental nursing class	116	29.1
	In total nursing education	62	15.6
	In hospital education	91	22.9

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Perception about the protection of medical information

Category	Min	Max	Mean	±SD
Privacy protection	1.00	5.00	3.37	.79
Confidentiality	2.38	5.00	4.20	.58
Medical information management	2.22	5.00	4.18	.64
Communication	1.86	5.00	4.20	.65
Total			3.99	.51

Regarding the perceptions for each question, Q6 (*"When presenting a case study in an academic meeting or to others after practice, a nursing student must not reveal a patient's personal information"*) had the highest score, at 4.50 ± 0.79, followed by Q25 (*"Nursing students must not give a patient's information to any person not related to the patient's treatment on the phone";* 4.48 ± 0.79) and Q24 ("*Nursing students must not disclose a patient's medical information obtained during practice in a private place"*; 4.46 ± 0.84).

In contrast, Q4 ("*When a nurse educates a patient about being discharged from the hospital, nursing students must prevent people other than the patient from listening to their conversation*") had the lowest score, or 2.97 ± 0.97.

3.4. Differences in Perception According to General Characteristics

Table 4 shows the differences in subjects' scores on the perception questionnaire according to their general characteristics.

In the categories of privacy protection and communication, perception score significantly differed according to school year. Senior students had significantly lower scores for privacy protection ($F = 4.107$, $p = .007$), but significantly higher scores for communication ($F = 2.725$, $p = .045$). In the case of having experience in hospital practice, 'privacy protection' (-2.61, $p = .010$) and medical information management (-2.05, $p = .041$) were statistically significantly lower. However, 'communication' (2.32, $p = .021$) was significantly higher.

The scores in the 'Communication' area were significantly higher when 'Medical Information Protection Act (3.40, $p = .001$)' and 'Nurse Code of Ethics (2.07, $p = .040$)' were known. In addition, when the subject completed 'medical security education' (-2.92, $p = .004$), the score in the 'communication' area was high.

4. DISCUSSION

Regarding the average scores of nursing university students' perceptions on the protection of patients' personal information, three of the four categories in the questionnaire had scores over 4.18 (out of 5). This result (4.11 point) is similar to that found in Lee's study (Lee & Park 2005), which was conducted with 289 nurses, but is lower than that found in (at 4.39) in Jung's (2011) study, which was conducted with nursing students using a 25-item questionnaire. Furthermore, the total mean score of the perception questionnaire was relatively high (at around 4 points). If a patient's personal medical information is

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Scores of questions on perceptions of protective actions concerning patients' medical information

	Questions	Scores	
		M	±SD
Privacy protection Protection	Q1 When a nurse checks a patient's health, the nursing student must prevent people other than the patient from listening to their conversation.	3.50	.93
	Q2 When medical staff talk about the patient's medical treatment in a hospital room, the nursing student must prevent people other than the patient from listening to their conversation.	3.38	1.06
	Q3 When a nurse tells a patient the results of an examination, the nursing student must prevent people other than the patient from listening to their conversation.	3.60	.98
	Q4 When a nurse educates a patient about being discharged from the hospital, the nursing student must prevent people other than the patient from listening to their conversation	2.97	.97
Confidentiality	Q5 To use a patient's medical information for research purposes, a nursing student must obtain the patient's agreement.	4.31	.94
	Q6 When presenting a case study in an academic meeting or to others after practice, a nursing student must not open a patient's personal information.	4.50	.79
	Q7* When a case study report is posted on the Internet, all types of information, including personal identity can be viewed.	4.35	1.12
	Q8 In order to record a patient's conversation, a nursing student must receive the patient's written agreement.	4.31	.86
	Q9 Nursing students must not tolerate a situation where a different colleague student records or photographs nursing care or treatment.	3.99	1.05
	Q10 Nursing students must not inform a patient's guardian of the patient's medical information without obtaining the patient's agreement.	3.76	1.06
	Q11 Nursing students must not disclose a patient's information without their agreement, even when the information is required for a lawsuit.	4.16	.92
	Q12 Nursing students must not talk about what they observe to people other than medical staff.	4.20	.89
Medical information management	Q13 When viewing data on a medical information system, nursing students must only view the minimal range of data directly related to the patient they are caring for.	3.85	1.04
	Q14 During practice, nursing students must not take a list of patients or medical records received by the nurse outside of the hospital ward.	4.24	.97
	Q15 Nursing students must help in preventing other people from looking at a patient's medical information.	4.29	.84
	Q16 To access the medical information system, nursing students must use their assigned ID or password.	4.36	.82
	Q17 Immediately after viewing data on the medical information system, the nursing student should log out.	4.37	.81
	Q18* Nursing students should share their own ID and password with others.	3.70	1.38
	Q19 Nursing students should not photograph medical information (EMR, PACS, or OCS) without receiving the written agreement of a patient.	4.23	.98
	Q20 Nursing students should not search for the medical information of their colleagues or celebrities for purposes other than actual practice.	4.33	.90
	Q21 Nursing students should reject any request of writing or modifying medical information untruthfully.	4.29	.93

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Continued

	Questions	Scores	
		M	±SD
Communication	Q22 Nursing students must not talk about a patient in an elevator or corridor.	4.34	.88
	Q23 When with other students, a nursing student must not talk about what they hear when medical staff are transferring patients' information between them.	4.06	1.06
	Q24 Nursing students must not disclose a patient's medical information obtained during practice in a private place.	4.46	.84
	Q25 Nursing students should not give a patient's information to any persons irrelevant to that patient's treatment on the phone.	4.48	.79
	Q26* Nursing students can give a patient's information to visitors of that patient.	3.61	1.36
	Q27 If requested to send a patient's information to an employee of another department that is not relevant to the patient's treatment, nursing students must not do so.	4.09	1.00
	Q28 Nursing students must not provide a patient's information to medical or insurance firms without the agreement of the patient.	4.39	.83

* Negative questions: Q7, Q18, and Q26

Table 4. Differences of the level of recognition in subjects characteristics

Category		Privacy Protection		Confidentiality		Medical Information Management		Communication	
		M ±SD	t/F (p)	M ±SD	t/F (p)	M ±SD	t/F (p)	M ±SD	t/F (p)
Grade	1	3.50 ±0.74[a]	4.107 (.007)	4.18 ±0.57	1.964 (0120)	4.21 ±0.72	1.561 (.199)	4.05 ±0.67[a]	2.725 (.045)
	2	3.52 ±0.65[a]		4.30 ±0.53		4.32 ±0.57		4.12 ±0.69	
	3	3.49 ±0.85[a]	a>b	4.04 ±0.64		4.14 ±0.67		4.18 ±0.75	a<b
	4	3.18 ±0.83[b]		4.21 ±0.59		4.12 ±0.61		4.32 ±0.54[b]	
Hospital practice experience	Yes	3.27 ±0.48	-2.61 (.010)	4.16 ±0.60	-1.26 (.210)	4.12 ±0.63	-2.05 (.041)	4.28 ±0.61	2.32 (.021)
	No	3.51 ±0.69		4.25 ±0.55		4.28 ±0.64		4.09 ±0.68	
Are you aware of the Medical Information Protection Act?	Yes	3.33 ±0.82	-1.02 (.310)	4.23 ±0.58	1.30 (.193)	4.20 ±0.62	0.72 (.473)	4.30 ±0.59	3.40 (.001)
	No	3.44 ±0.73		4.13 ±0.59		4.15 ±0.67		4.01 ±0.70	
Are you aware of the Korean Code of Nursing Ethics?	Yes	3.35 ±0.81	-.65 (.516)	4.21 ±0.58	.90 (.368)	4.20 ±0.62	.75 (.453)	4.24 ±0.64	2.07 (.040)
	No	3.44 ±0.72		4.13 ±0.59		4.12 ±0.71		4.02 ±0.67	
Received education on medical information security	Yes	3.31 ±0.78	1.56 (.120)	4.21 ±0.60	-0.39 (.696)	4.21 ±0.62	-0.72 (.471)	4.30 ±0.58	-2.92 (.004)
	No	3.46 ±0.81		4.18 ±056		4.15 ±0.67		4.06 ±0.71	

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

leaked, it is highly likely to violate that patient's privacy. Therefore, it is essential, rather than optional, to protect patients' medical information. Given that, when planning an educational program, educational institutions must take into account the scores of each category found in this study and provide tailored education that contributes to accomplishing a final goal (5 points). Fortunately, medical institutions involved in nursing students' clinical practice often make some effort to protect patients' personal information (Jeong, 2016).

With regard to nursing students' perceptions by category, the communication category had the highest score. This result is similar to that of a previous study [7, 8, 9]. The high score in the communication category is perhaps because it includes questions relating to a high risk of personal information leakage, including Q24 ("Nursing students should not disclose a patient's medical information obtained during practice in a private place"), Q22 ("Nursing students should not talk about a patient in an elevator or corridor"), and Q23 ("When with other students, a nursing student should not talk about what they hear when medical staff are transferring patients' information between them"). Additionally, it might be because students about to engage in clinical practice receive intensive education on this topic at nursing school. Therefore, the repeated education seems to bring about its effect.

The differences in perceptions were analyzed according to study subjects' characteristics. The results showed that junior and senior students tended to have higher perception scores than did sophomore students. This is perhaps because, with every subsequent school year, students are likely to receive repeated education about patients' personal information protection, either at school or during their clinical practice at a medical institution, which in turn influences their perception. Particularly, subjects who had previous experience of clinical practice, who were aware of the Medical Information Protection Act and Korean Code of Nursing Ethics, and who had received education on medical security had significantly higher scores in the communication category compared with other categories. Therefore, this study proposes that schools and medical institutions should attempt to discuss effective education strategies in consideration of their own educational environments and develop a systematic educational program that improves students' perception of protecting medical information.

5. CONCLUSION

In order to improve the perception of nursing students about the protection of patients' personal information, it is necessary for schools and medical institutions to establish educational strategies in accordance with their environment. It is also important to realize the importance of personal information, to develop sustainable and systematic education programs.

REFERENCES

Jeong, H. C. (2016). An analysis on nursing students' perception of protective actions and management for medical information at hospital. In *Proceedings of International workshop of Healthcare and Nursing 2016*, Jeju, December 19-21. 10.21742/ijria.2016.4.2.04

Jeun, Y. J. (2013). EMR System and Patient Medical Information Protection. *The Korean Journal of Health Service Management*, *7*(3), 213–224. doi:10.12811/kshsm.2013.7.3.213

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jin, H. E., & Choi, E. M. (2012). A Study on Factors Affecting the Reception Attitude toward Electronic Medical Record. *Journal of Digital Convergence*, *10*(4), 279–286.

Jung, K. I., & Jung, H. Y. (2011). A study on the student nurse's perception and performance protecting behavior for the patient medical information. *The Korean Journal of Health Service Management*, *5*(4), 65–79. doi:10.12811/kshsm.2011.5.4.065

Kim, C. H., Jeong, S. Y., & Song, Y. S. (2013). Recognition and Performance of Patient Private Information Protection (PPIP) in Nursing Students. *Journal of Digital Convergence*, *11*(11), 479–490. doi:10.14400/JDPM.2013.11.11.479

Kim, H. E., & Kim, J. H. (1990). A Survey on the attitude of social groups toward security. privacy, and confidentiality of health information: an original paper authors and affiliations. *Journal of Korean Social Medical Information.*, *5*(3), 63–76.

Koskosas, I. V., & Asimopoulos, N. (2011). Information System Security Goals. *International Journal of Advanced Science and Technology*, *27*, 15–26.

Lee, M. Y., & Park, Y. I. (2005). A study on the nurse's perception and performance of protecting patient privacy. *Clinical Nursing Research*, *11*(1), 7–20.

Schick, I. C. (1998). Protecting patient's privacy; health information networks raise new questions. *Health Progression*, *79*(3), 26–31. PMID:10180017

Suh, C. J., & Lee, I. T. (2014). The effects of EMR on process improvement and operation performance. *Journal of the Korea Service Management Society*, *15*(3), 327–345. doi:10.15706/jksms.2014.15.3.013

This research was previously published in the International Journal of E-Health and Medical Communications (IJEHMC), 9(4); edited by Joel J.P.C. Rodrigues; pages 11-19; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 84
Information Security Standards in Healthcare Activities

José Gaivéo
Polytechnic Institute of Setubal, Portugal

ABSTRACT

Information is mandatory in healthcare activities and in all that are related to it. In this same sense, people that deal with those information requires attention because patient´s information could be exposed. The use of directions stated by information security standards might allow a proactive attitude in the face of the diversity of threats that as the potential to explore the vulnerabilities of organizational assets. This article intends to recognize information threats and vulnerabilities that could be explored, using information security international standards to support the activities needed to assume information safeguard. Another intention is the establishment of a basis of references in information security to define a level of risk classification to build a referential to the potential that a given threat has to exploit the vulnerabilities of informational assets, preventing damages to personal and organizational property, and also activity continuity, assuming information as the main resource.

INTRODUCTION

Nowadays, information, one of the great challenges of societies, implies profound changes in the vast majority of institutions, not only in its functioning but also in people.

The quantity and diversity of information that is currently available and easily accessible by the emergence of an information society brings new and hard challenges. This type of society deals essentially with information and need people with appropriate skills and competencies to deal with it, using all available resources to assure that information is adequately used and, consequently, protected.

Facing the current context of social and economic globalization, in which a society based on the use of information emerges, it is important to use appropriate technologies and systems to serve institutions. This information systems and information and communication technologies (ICT) support can and should be used to conduct institutional development based on information.

DOI: 10.4018/978-1-7998-1204-3.ch084

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As can be noted, information has seen growth in its importance within the institutions, being essential ensuring its widespread use. In this sense, the institutions, given the challenges and implications that arise from the modernization necessary to their development, require that its employees, experts or not, have the right skills for the use of information systems and ICT to support information needs.

Thus, security of information systems and ICTs is only a part of information security in an organization. Of course, the security of information systems and ICTs is important in the organizational context, mainly because information systems and ICTs are pertinent to support information acquisition, supporting, storing and sharing according to organization needs.

Nevertheless, taking simply information systems and ICT into consideration, information safeguard covers only a part of the whole problem, what implies that the approach to the security of organizational assets must be focused on whole organization, which means that security measures should attend to it.

In this sense, the people awareness and training about security are subjects that must be taken into account. This is so relevant as the potential impact of information on information systems and ICT usage, and also in the activities related to information about patient's health.

Because patient's data must remain confidential, organizations shall adopt information security policies to provide the appropriate protection. Those policies could use directives provided by information security international standards as ISO/IEC 27x family, in order to establish procedures and rules that correspond to organizational business needs.

With this paper, through the assumption that information is an important organizational asset, and that is of particular importance in the healthcare sector, is intended to make the appropriate recommendations for the establishment of a framework to support the establishment and implementation of information security policies that provide the adequate protection of information.

In this circumstance, also, by the nature of the asset information, consider the role of people who deal with this information and provide them with the necessary awareness of this problem, as well as with the education and training suited to each concrete context.

We must also, in addition to the identification and valuation of such assets, identify threats and vulnerabilities that can affect the assets, in order to assess the risk and, consequently, to its management.

Still intend to characterize what is meant by information security and their supports, getting to know the legal and regulatory framework to ensure compliance of security measures, with international information security standards, such as ISO/IEC 27x family, to support the main activities related to these objectives.

BACKGROUND

In a sector where main activities are conducted by people, as healthcare, the foremost resource relates to information about patient's life. In this sense, information needs to be adequately protected against misuse and disclosure of any kind.

Few years ago, health patient's data were essentially supported by paper, and the major problems was their handling and storage, this contributes to several difficulties in its usage. Nowadays, despite people are more awarded about information systems and ICT usage implications, data and information remain without appropriate attention related to their security.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Many of these challenges involve structural changes in organizations as a factor of sustainability of the information society in the current context of globalization of markets and societies. Those changes can, and should, involve profound changes in the structure of institutions requesting membership, and the conscious and active participation of all stakeholders.

However, it should be observed, those information supports cannot be assumed as a panacea for all problems in institutions, but must be understood in its essence as enhancers of competencies of people and its capacity for the use for the benefit of institutions.

On the other hand, this new attitude embodies a clear perception of the relevance of information to the organization and the assumption of the need to provide their protection according to their criticality in line with the objectives and strategy. Protect information and systems that support their sharing internally and with the various partners emerges as a key strategic objective, in the context where the credibility of the institutions is a critical success factor in their activity.

Think about information security implies, besides to think about information as an organizational asset, to think about the security of their main supports in today organizations, namely information systems and ICTs. This attitude that must be assumed as critical to business success.

Thinking about how to solve those simple problems, allows us to note that potential solutions aren't as easier like initially seeming to be. Mainly because there are two main issues that could affect any possible solution; a) people, due the ways they use information, and b) the information themselves. Those issues are factors that tend to increase the vulnerabilities, and consequently the threats that could explore those vulnerabilities.

Organizational Assets and Security

Organizations are composed of people, and the technologies can help people to perform their activities. However, the use of technologies as competitive advantage depends on how they are used, and also how people and information are understood as organizational resources.

Sharing information can affect power relations within the organization, implying that the success of the information systems depends of organizational decision-makers, not excluding, however, that also depend on end-users.

Being the information a resource, it is necessary to ensure their possession, where the protection of information requires the availability of adequate resources, a fact which by itself needs more resources, implying an increase in costs.

Information is usually defined as something that has value to an organization and that requires adequate protection, regardless of the form or medium of storing, processing, and/or transmission, but could be also defined as a business asset that is essential to an organization and consequently needs to be suitably protected (International Organization for Standardization, 2014). This vision is reinforced by Workman, Phelps, and Gathegi (2013) that presents information as data interpreted in a specific context.

However, in a sector like healthcare, information is not only about patients, but also about all participants i.e. include, besides patients, collaborators, stakeholders, competitors and regulators.

One of the aspects that highlights to the need for information security, today is the greater availability of information and the ways of use. this has increased problems brought to organizations and people, making the need for information security in a mandatory factor; especially in organizations where information is an asset essential to business.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nevertheless, in a general definition, information security is seen as the preservation of confidentiality, integrity and availability of information, something that Whitman and Mattord (2010) mentions as "the key concepts of information security, which are mainly referred as CIA triangle" (p.8).

In this position, confidentiality is to ensure that information is only accessed by those authorized to do so. Integrity is about not modifying the information ensuring its accuracy and credibility. Availability is to ensure that authorized persons have access to information and associated services whenever necessary (Smith, 2013). To do this, according to Whitman and Mattord (2010); confidentiality implies, among other measures, to classify information, security documentation and apply security policies, while integrity involves to take care about information entering, storing and sharing, and finally, availability relates with the information format that allow their usage without interference or obstruction.

Complementarily, Smith (2013) assumes that security process has six phases: identify assets; analyze the risk of attacks; establish security policies; implement defenses; monitor defenses; and, recover from attacks. The author also notes that "defense is a security measure intended to protect the assets from the threat agents" (p.12).

The realization of that objective is achieved through the definition and implementation of information security policies that must provide the management's commitment and appropriate support of critical activities to information security, where security policy is the set of good practices that the organization must follow and thus the expression of organizational strategy.

Information security develops through an information security management system (ISMS), which includes the collation of business assets (people, equipment, buildings, auxiliary material and information, among others) that are identified as essential for the normal development of the business by providing their management in accordance with the stipulations of the standards and policies existing in the organization.

The ISMS is a part of the general system for the management (includes organizational structure, policies, planned activities, responsibilities, practices, procedures, processes and resources), based on the approach to business risk, to establish, implement, operate, monitor, review, maintain, and improve an organization information security to achieve business objectives (International Organization for Standardization, 2014).

According these points of view, it's important to know and understand legal and regulatory framework to adequately develop an information security strategy that looks to information protection in healthcare sector. In this subject matter, because information is related with patients, it's important to know who can see and use that information, in what circumstances and what for.

Legal Framework

In terms of supporting legislation, it is transverse in most countries the existence of laws for the protection of personal data. Some examples are:

- Personal Information Protection and Electronic Documents Act of 2000 (PIPEDA) – Canada;
- Federal Information Security Management Act of 2002 (FISMA) – United States of America;
- Data Protection Act 1998 – United Kingdom;
- Data Protection Act of 1978 (revised in 2004) – France;
- Data Protection Directive of 1998 – European Union;
- Federal Data Protection Act of 2001 – Germany.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In a general way, countries adopt several laws and regulations that complement those and allow states, institutions, enterprises and people to perform tasks to provide information security.

Some of these countries understand the importance and sensitivity of the information on this sector in particular and adopted specific laws/regulations to provide a more objective and greater protection of that information. In this context must be referred the Health Insurance Portability and Accountability Act (HIPAA, Public Law 104-191), which was approved by the Senate and House of Representatives of the United States of America, by 1996.

Later, in 2003, related with this same law, was approved a security rule (referred to as HIPAA Security Rule) that constitute the security standards for a reliable information on health (Scholl *et al*, 2008).

In addition to ensure the portability of health insurance, HIPAA sets a number of requirements relating to security and privacy of patient information (HIPAA, 1996).

Although this seems only relevant to the health care industry, any organization that transmits and store patients' health information, including information on employee benefits and claims, is subject to the rules of HIPAA.

The main purpose of this Rule is to ensure that the health information of anyone is adequately protected, while the flow of health information required to a) provide and promote healthcare and b) high quality to protect public health and welfare is also protected.

To do so, Whitman and Mattord (2010) reminds the five fundamental principles of HIPAA, which are:

- Control of medical information by patients;
- Use medical information with boundaries;
- Privacy of private information accountable;
- Balance responsibility for the use of medical information;
- Security of health information.

This rule searches for a compromise that allows important uses of information, while protecting the privacy of individuals seeking health care. The design of the principles must obey to the indispensability of being flexible and comprehensible in order to cover a variety of uses and discoveries that need to be addressed.

To achieving that, HIPAA restricts the access to private health information without documented consent, giving patients the right to know who as access to their information.

Countries have not their information security regulation in the healthcare sector, being frequent the adoption of several HIPAA procedures to sustain most of their needs in this area.

Regulatory Framework

Although there are many rules about information security, which fall into a wide range of aspects associated with implementing information security policies and support systems, some are part of internationally recognized standards such as ISO/IEC 27x series, which assumes the security as a management process by providing framework guidelines to safeguard organizational information.

Because the ISO/IEC 27x series standard assumes the information security issues as a management process, it is important to make use of a common and consensual definition of information security provided by these standards as the preservation of the confidentiality, integrity, and availability of information (International Organization for Standardization, 2014).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In this subject matter an international standard on security area as ISO/IEC 27002, known as the code of best practices for information security, provides some guidelines for controls related to mobility of information and with devices in transit throughout the organization.

Those guidelines of controls, used to sustain risk assessment performed according to ISO/IEC 27001 through the application of ISO/IEC 27005 or ISO 31000 risk management models, allow the implementation of security measures to protect information assets and their supports.

The previous reality reinforces the relevance of a standard like ISO 27799, which applies ISO/IEC 27002 concepts, to be used as the main base to provide a key reference to deal with the multiplicity of security issues that emerges from an enlargement of the amplitude provided by information on social and organizational activities in healthcare sector.

Usually, ISO 27799 applies their controls additionally to those provided by ISO/IEC 27002, attending to the specificity of issues related with healthcare matters, i.e., the guidance provided by ISO/IEC 27002 must be considered in addition to the guidance provided by ISO 27799.

Using a standard like ISO/IEC 27001: 2013 (Information security management systems – Requirements) helps to establish an ISMS that intends to support the protection of the organizational information and other business assets that are fundamental to organizational activities.

The ISO/IEC 27001 specifies requirements for establishing, implementing, maintaining and improving the ISMS in the context of the business of the organization (International Organization for Standardization, 2013a).

The standard is organized into seven distinct but complementary chapters (International Organization for Standardization, 2013a), has the following:

- Context of organization;
- Leadership;
- Planning;
- Support;
- Operation;
- Performance evaluation;
- Improvement.

Based on those chapters, the ISO/IEC 27001 provides a set of steps to assure a security program:

- Set the ISMS and its scope;
- Set the information security policy;
- Define roles and responsibilities;
- Develop an appreciation of the security risks;
- Manage the identification of risks;
- Select the controls to be implemented;
- Prepare the appropriate documentation, including a statement of applicability;
- Provide information security motorization and evaluation;
- Establish the needed improvement;
- Assure compliance.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The structure of this standard has in attention an implementation approach that could be understood as being based on PDCA (Plan-Do-Check-Act), also known as *The Deming Cycle*, comprising the following aspects:

- Plan - Establish the ISMS;
- Do - Implement and Operate the ISMS;
- Check - Monitor and Review the ISMS;
- Act - Maintain and Improve the ISMS.

The ISO/IEC 27002:2013 (Code of practice for information security controls) presents the set of good practices that organizations must follow to perform information security management, providing a list of controls that may be used in the framework of ISO/IEC 27001.

This standard has 114 security controls framed in 14 security control clauses (International Organization for Standardization, 2013b); each one organized in main security categories that group the controls, which is as follow:

- Information security policies;
- Organization of information security;
- Human Resource security;
- Asset management;
- Access control;
- Cryptography;
- Physical and environmental security;
- Operations security;
- Communications security;
- Systems acquisition, development and maintenance;
- Supplier relationships;
- Information security incident management;
- Information security aspects of business continuity management;
- Compliance.

The order that the clauses are presented in this standard has nothing to do with its importance, but with the circumstances of the application according to organizational security needs, leading to every organization that uses this standard can identify the applicable clauses their relevance and application to individual business processes.

The primary purpose of information security is, in accordance with this standard, to protect the information from a variety of threats, ensuring business continuity through the minimization of damage to the business for security incidents, seeking to maximize the return on investments and business opportunities.

This cycle includes the assessment of risks, implement controls for dealing with risks, monitor and review the risks, and maintain and improve the system of risk controls.

Harmonized with the stipulations in the main reference standards with regard to information security, in particular ISO/IEC 27002 and ISO/IEC 27001, the ISO/IEC 27005: 2011 (Information Security Risk Management) presents a set of guidelines that aim to enable a more efficient risk management in organizations, considering all aspects of risk management cycle of an ISMS. In the same topic, ISO 31000:2009

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

(Risk management) presents an approach that provides a wider risk assessment due to a focus to whole organization. Nevertheless, assuming that both standards could be used, ISO/IEC 27001expressly refers that the information security risk assessment and treatment processes must follow ISO 31000 principles and generic guidelines, however, this recommendation didn't imply the exclusivity of use.

In this sense, the ISO/IEC 27005 was especially developed to be used as a tool for business managers and information security professionals directly or indirectly involved with the development, implementation, operation, management, maintenance, revision and/or organizational continuous improvement of the ISMS, and should be implemented according to an approach by processes (International Organization for Standardization, 2011).

The structure of ISO/IEC 27005 is divided into 4 groups of guidelines that allowing in particular:

- The framework of the concept of risk to information security on an organizational context;
- The process of analysis and assessment of risk in the organization;
- The process of risk treatment and management, also as regards decision-making on the part of organizational management;
- Continued activities for risk management within the organization's ISMS.

While ISO/IEC 27005 is focused on risk management to information security, ISO 31000 relates to risk management to organization and provides principles and generic guidelines that can be applied to the complete lifecycle of an organization.

The structure of ISO 31000 is divided into two major directives: Framework and Processes. The framework focuses on the risk management and on organizational context, while Processes focuses on risk assessment and risk treatment and monitoring (International Organization for Standardization, 2009).

This standard presents a framework for risk management that uses the same principles of PDCA cycle to design, implement, monitor, review and continual improvement of the framework.

By 2008 was published the first edition of ISO 27799 (Health informatics – Information security management in health using ISO/IEC 27002), that provides guidance to healthcare organizations and other custodians of personal health information, on how to best protect the confidentiality, integrity and availability of such information by implementing ISO/IEC 27002 (International Organization for Standardization, 2008). Due to the extensive application of ISO/IEC 27002 for health informatics security management through several countries, namely Australia, Canada, France, United Kingdom, etc. this standard, the ISO 27799, complements the more generic standards as ISO/IEC 27002 or ISO/IEC 27001, applying, besides ISO/IEC 27002 controls, the PDCA model that supersedes those family of standards related with information security to continuously maintain the information security program.

This statement allows the application of same perspectives and studies, the majority of them based in ISO/IEC 27002 for instance, to information, and also to their supports, related with healthcare activities. The application of ISO/IEC 27002 controls – framed by ISO 27799, due its relevance and implantation in most organizations, constitutes an added value factor to a sector, like healthcare, where information protection isn't a major preoccupation, but something that must be done.

The standard ISO 27799 (2008) has the following chapters:

- Health information security;
- Practical action plan for implementing ISO/IEC 27002;
- Healthcare implications of ISO/IEC 27002.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

However, this standard provides a set of steps to implement a health information security program:

- Management commitment to implementing ISO/IEC 27002;
- Establish the ISMS (phase Plan of PDCA model);
- Implement and operate the ISMS (phase Do of PDCA model);
- Monitor and review the ISMS (phase Check of PDCA model);
- Maintain and improve the ISMS (phase Act of PDCA model);
- Appliance of guidance found in ISO/IEC 27002 to healthcare issues.

Nowadays, the actual version of this standard, dated from 2008, is under a review process to align with the new versions of ISO/IEC 27001 and 27002, published in 2013, with the expectations that will be published in 2016. In this sense, should be noted that there aren't expectations of huge changes between versions that imply substantial alterations in the practices to be adopted within healthcare sector related to information security.

In addition, should be considered the suitability of incorporating the recommendations made by the standard ISO/IEC 27005 (International Organization for Standardization, 2008a), particularly when the ISO 27799, in annex A, takes into account aspects relating to threats to the confidentiality, integrity and availability of health information assets, enabling an objective approach to the treatment of risk regarding information security.

The issues referenced by annex A of ISO 27799, in articulation with ISO/IEC 27002, are: Masquerade by insiders; Masquerade by service providers; Masquerade by outsiders; Unauthorized use of a health information application; Introduction of damaging or disruptive software; Misuse of system resources; Communications infiltration; Communications interception; Repudiation; Connection failure; Embedding of malicious code; Accidental misrouting; Technical failure of the host, storage facility or network infrastructure; Environmental support failure; System or network software failure; Application software failure; Operator error; Maintenance error; User error; Staff shortage; Theft by insiders; Theft by outsiders; Willful damage by insiders; Willful damage by outsiders; Terrorism.

INFORMATION SECURITY ISSUES

The problems inherent in the management of data (sorting, handling, storage, availability) and the management of information (production, sharing, and utilization) must, given the importance they possess now, be a constant organizational decision-makers concern.

In this matters, could be followed some recommendations made by Gaiveo (2011) that refer the importance of a proactive approach to information security requiring, from managers and people, new attitudes and a larger mutual aid that goes further than what were the usual day to day activities, particularly using the provided identification of ISO/IEC 27002 controls related with people and with information supports, which could be applied according to organization needs.

Security-related issues have increased relevance when there are large volumes of data or a large number of potential users, particularly when the shapes and tools for accessing these data are diverse and available.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Breaks in security arise habitually following the occurrence of disasters that may manifest themselves in various forms; may be natural (earthquakes, fires, floods, lightning) or unnatural (technological, manifesting itself in failures of systems, and human, appearing distinctly as undue access, theft, and/or sabotage).

The success of security measures depends on the organizational commitment of decision-makers and the involvement of people in pursuing its objectives, particularly when there is an information society where organizations depend on their information to fulfill their business strategies.

It should, however, be noted that the optimal solution to a problem of this kind must pass by a security policy that involves the whole organization in all its aspects. This does not preclude, that there may exist, it is even desirable, a classification on the basis of the levels of the criticality of the organizational information.

This way, and articulate with the legislation which provides the accounting framework for public administration and the diverse normative applicable, it is important to carry out a regulatory framework for information security. That framework must seek to establish the appropriate mechanisms for safe and effective use of accounting information arising from the activity of the institutions. This could be done by assuming reference standards that are able for performances of excellence to the development of an information society more inclusive in all aspects, involving the public and private sectors in a partnership of national development.

Besides to the usual threats and vulnerabilities (and consequently risks) that any information is subject, this type of information is often the target of greatest attention, requiring the strengthening of preventive measures and a greater awareness of their protection.

In this sense must, therefore, be established a set of aspects to consider as deserving of special attention, namely: people, vulnerabilities, threats, attacks, risks.

From this, emerges the need of understanding security concepts like vulnerabilities, threats, and risk, as well the people involvement and responsibilities in security matters, and even security requirements to support information security policies definition according to specific needs of the healthcare sector.

Involvement and Responsibility of Individuals

Nowadays, the main challenges are related to people and with the ways that they act in societies, organizations, considering personal, social or institutional relations. Moreover, considering information as the main asset of organizations elicits the need to strengthen the protection of information and consequently assure information security according to legal and normative framing, in order to provide their appropriate usage to support business activities.

In the current social context, all organizational functions play an essential role in the activities of business, no matter if are functions directly related to the business, or if are supporting functions only.

Compete with the demands of new markets and customers is not possible without the involvement of all organizational levels as a unified structure that should function in an integrated and coherent way.

The information has come to impose its presence on all organizational levels, becoming a feature of extreme relevance for most of the organizational functions.

Commonly referred as the weakest link in matters affecting the security of information systems and technologies that support the flow of information, people represent an important organizational resource.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

How people react to change their behavior to using ICT, particularly those affecting their everyday activities, largely depends on the dissatisfaction at several levels, whether for reasons of adaptation and organizational culture or even incompetence, conducting to not acceptance of the rules.

In this context, financial and accounting functions use information which, by its relevance and criticality, must be protected to ensure its confidentiality, integrity, and availability. This observation implies the existence of an information security policy that is appropriate to ensure the privacy of such information.

In an organization, successful information security results from a balance between three distinct issues: organization objectives; risks; and, security measures costs (Smith, 2013).

Another aspect to consider relatively to the involvement of people is linked to the so-called "social engineering," that combines technical skills with the attitude of pretending to be someone that the attacker is not to harm or damage organizational assets (Smith, 2013).

The involvement of people in institutional activities soon reveals the importance of its appropriate framework for these activities and their attitudes towards the information they use. These visions are corroborated by Whitman and Mattord (2010), to whom "people are the most critical link in the information security program" (p.15), what allows the assumption that is crucial the raising of awareness of information security and potential organizational impacts arising from these situations.

On the other hand, it is equally important to provide adequate training on the organizational strand, looking for a greater involvement of people in the use of organizational resources to enhance their performance.

The pursuit of these objectives relating to the persons, as for the most of the activities related to organizational security, obliged to a clear and unambiguous commitment of top management, thus making clear an involvement at all organizational levels in the activities essential to the development of institutional activities.

The involvement of people in these activities is crucial to the success of the safety objectives defined for the organization mainly due to the fact that, in the contexts referenced, people and information have become fundamental assets in business organizations by allowing them to respond adequately and timely to the new requirements of the information society.

Then, according to Dark (2006), considering people as the first line of defense in terms of security, organizations need to provide awareness raising, education, and training in this area.

The role of people assumes new scopes, including the use of information and guarantee information security of business, boosting its full use by the organization in accordance with the needs identified in the conduct of its business activities.

This is so much serious when attacks come from internal sources, mainly because inside attackers have special access to valuable assets, i.e. they have almost direct access to the resources needed to perpetrate the attacks (Smith, 2013).

The need to ensure the confidentiality and reliability of the information were always constant since the dawn of humanity, which varied according to the way of sharing such information (including its acquisition and treatment).

The universe of general sharing of the information available was initially very low (the family). With the human evolution, gradually extending to the present day, the sharing of information is quite unlimited by the access to existing technologies and the capacity for its use.

Emerges thus strengthened the indispensability of assume people's relevance in the context of information security, and simultaneously encourage a greater awareness of this type of problem. Because

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

people are frequently referred as the weakest link in any security matter, is crucial that, according to Merkow and Breithaupt (2014) "that a security program address user education, awareness, and training on policies and procedures that affect them" (p.72).

Vulnerabilities, Threats and Security Risks

There is often confusion between fundamental concepts in information security (vulnerabilities, threats, risks) that matter to specify and to frame.

Organizations face some difficulties when trying to define what information should be protected and what level this protection must achieve. Is the definition of the risk associated with the likelihood and potential damage that allows support options necessary for the framing of the definition of security policies and ISMS.

Vulnerability can be seen as a weakness or flaw in a system that exposes information assets to attack or damage (Whitman & Mattord, 2011), or, according to Smith (2013), "as a weakness in the boundary that protects the assets from the threat agents" (p.12), which allows to assume that vulnerabilities are weaknesses in the systems, which can be exploited by the threat materializing. In the same sense, Workman *et al.* (2013) assume that "vulnerabilities represent the potential to exploit systems and information" (p.55).

As examples of vulnerabilities can refer:

- **Related to People:** Absences; death; insanity; competencies; only one; ethics; errors; sensitivity; communication;
- **Related to Equipment:** Malfunction; power failures; portability; depreciation;
- **Related to Information:** Portability; format; limited life; single; legal compliance.

Threats are events that can lead to potential damage and cause undesirable effects

The source of threats could be internal or external and are linked to the vulnerabilities of a system, what means that the exposures of those vulnerabilities are strongly linked with the threatening agents (Workman *et al.*, 2013).

Whereas the likelihood that a given threat exploiting a given vulnerability successfully is estimated by the opportunity and motivation of the source of the threat or by methods which lead to this exploitation, while the impact of a successful exploitation is estimated by examining the effects of exploitation may have on confidentiality, integrity and availability of the system and the data it processes (Bowen *et al.*, 2006).

Risk may also be defined as the potential of a threat to harm or damage people or assets (Workman *et al.*, 2013), as the negative impact of performance on a vulnerability considering the likelihood and the impact this occurrence (Stoneburner *et al.*, 2002), or as the probability that a particular threat to exploit a particular vulnerability (Peltier, 2010), or even the potential situation that threat the assets (Smith, 2013).

Risk can then be defined as potential losses incurred by a business, as a result of an unwanted event that can turn into losses in business or even interruption.

Risk, in a formal sense, corresponds to "danger; the possibility of being in danger", however, in terms of business, has two aspects that must be specified. The first part is directly related to the business and essentially depends on customers and competitors. The other strand relates to business support tools and activities that are related specifically to information, i.e. with its safety in relation to the risks which could affect it.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The difficulty to fully understand the concept of risk is inherent in the uncertainty, and, according to Workman *et al.* (2013) occurs when a known or unknown vulnerability is associated to a corresponding threat.

In risk determination, Merkow and Breithaupt (2014) consider that are two main factors to attend, being firstly the consequences of losing an asset and secondly the likelihood of that occurrence. According to the same authors, both factors must be identified and graduated, also giving an example (p.26) that frames likelihood between "rare" and "almost certain", and consequences between "insignificant" and "catastrophic."

Usually, the risk is the product of three factors: Threat, Vulnerability, and Value of the asset. In this sense, Merkow and Breithaupt (2014) note that "a risk is real when there is a presence of a threat, a vulnerability that the attacker can exploit, and a high likelihood that the attacker will carry out the threat" (p.72).

Risk management is defined as coordinated activities to direct and control an organization with concern to risk (International Organization for Standardization, 2009), which is achieved through the application of management policies, procedures, and practices.

The management of risk can be understood as the main security management issue (Workman *et al.*, 2013), in which context the risk management process is the assessment of information security risks, being this appreciation required to understand the security requirements of the business information, and risks to the organization's business assets.

The appreciation of the risk must identify, quantify and prioritize risks against acceptance criteria of risk and the relevant objectives for the organization, and the results obtained used to guide and determine priorities and appropriate management actions for the management of information security risks as also to support the implementation of selected controls that provide protection against these risks (International Organization for Standardization, 2013b). It is still referred the need of the valuation of risk that includes a systematic approach to estimating the magnitude of risk (risk analysis), and the process of comparing the risk estimated by the criterion of risk, to determine the relevance of risk (risk assessment).

In this context, Workman *et al.* (2013) refers "risk assessment as an ongoing process of identifying risks and threats" and risk management as "the ongoing process of implementing measures according to the costs and benefits associated with risks and security countermeasures" (p.105) incorporating also strategies and measures to recover from attacks.

From Smith (2013) perspective, risk assessment process has two separate activities, one corresponding to the identification of risks, i.e., how they arise and attack the assets, and the other related with the prioritization of risk by looking for the likelihood of occurrence and impact on the assets. To the author, those activities are performed through "five steps: Identify assets; Identify threat agents and attacks; Estimate the likelihood of attacks; Estimate the impact of attacks; and, Calculate the relative significance of attacks" (p.19).

According to ISO/IEC 27002 (International Organization for Standardization, 2013b), the valuation of risk (or risk assessment) includes a diverse set of activities and actions, among which are:

- Identification of the assets;
- Identification of legal and business requirements that are relevant to the identification of the assets;
- Valuation of assets identified, taking into account the legal and business requirements that are identified, and the impact of the loss of confidentiality, integrity and availability;
- Identification of vulnerabilities and significant threats to the assets identified;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Evaluation of the likelihood of occurrence of a vulnerability or threat;
- Calculation of the risk, namely its potential impact estimation.

The risk management process will comprise a number of aspects to monitor, in particular (International Organization for Standardization, 2013b):

- Avoid the risk;
- Transfer the risk;
- Reduce the vulnerabilities;
- Reduce the threats;
- Reduce possible risks;
- Detect unexpected events, react and recover;
- Accept the risk (residual).

Relatively to the processing of risk, organizations must decide in advance the criteria for determining when the risks are acceptable or not (International Organization for Standardization, 2013b). After that, for each of the identified risks following the appreciation of the risk, it is necessary to make decisions about the treatment of risk.

As a component of risk management, risk mitigation starts, according to Workman *et al.* (2013), with the identification of assets and this valuation. Those assets include systems software and information that must be evaluated to reflect their relevance to business activities.

Identification and Specification of Security Requirements

Acquire, select, store and share data depends mainly from the technological capabilities available and usable by the organization and also from people's capacities to use organizational resources to manage information. In this sense, produce, manage and share information depends crucially on the adequacy of the available technologies and human resources capacity in their use for the benefit of the organization. Already the creation and sharing of knowledge depend on the skills of human resources at the service of the organization.

If the development of new technologies facilitates the access to information, enabling a broader diffusion by all business partners, people involvement is a critical success factor to increase the accomplishment of information usage in business activities.

However, the success of information sharing is directly associated with the capacity of information support to respond to the organization's needs.

Information flows, supported by information systems, traverse not only the value-chain of the organization but also those of its partners, imposing the adequate skills and competencies to deal with it in order to increase the creation of value for all stakeholders that are involved. Thus, sharing information is not only within organizations but also among business partners.

Have data to produce information relevant to support activity and provide this information when it is needed and in the proper format, requires technological support, but also human resources with the necessary skills and knowledge.

The identification of security needs depends firstly on the intrinsic characteristics of each organization, but also of the resources involved, in particular of persons in its various organizational levels.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

This situation implies, however, identifying and classifying assets to protect, so as to enable the measurement of the effectiveness of options set to ensure the protection of these assets, and take the necessary measures to avoid or minimize potential threats to which they may become subject.

A clear specification of fundamental assets for organizational activities constitutes the basis for the establishment of adequate security policies to the reality of each organization.

Showing a perspective in the context of the necessary resources for developing and implementing security policies, Weaver, Weaver, and Farwood (2013) classify assets into four types:

- **Physical:** Equipment and buildings of the organization;
- **Data:** Databases, personal records, customer information and other data stored;
- **Application Software:** Server programs, security programs and other applications used in communications of organizational activities;
- **Personal:** People who work in your organization, customers, business partners and subcontractors that contribute to the organization.

Whereas this engaging, Peltier (2010) refer that, in order to be effective, the analysis process should be accepted as an integral part of the business process of the organization, stating that the risk analysis should seek the organization's business needs and establish safeguards to address these needs.

The identification of information assets to protect, the specification and characterization of the vulnerabilities that those assets are exposed, as well as also the identification of potential threats that could exploit those vulnerabilities, are fundamental aspects of an assessment of the risks that such assets are subject.

In this situation, the selection and characterization of the relative weights of each component is an essential factor for the adequate perception of the organizations regarding the relevance and impact of each threat.

Any approach to information security needs to deal with uncertain, and the culture and organizational environment, situation that leads to a real understanding of the particularities of each institution, mainly because: asset identification and classification, identification of threats/vulnerabilities, and risk assessment are specific, i.e. there have not an identical solution for all, by that, these activities shall be made according to the specificity of each organization.

That assumption reinforces the need of having guidance to fulfill information security requirements, essentially when ISO 27799 identifies threats to health information assets but didn't present asset classification, vulnerabilities likelihood, neither risks assessment. According to those issues; this is a situation that makes difficult their full applicability to a large part of information security policies.

A possible solution is provided by ISO/IEC 27005 that shows same examples that could be applied to information security in the healthcare sector, providing a risk assessment that is a crucial aspect in the context of information security, enabling a proper guiding of the resources given to assets that must be protected.

However, also, ISO 31000 could be applied to provide a must wide-ranging approach to risk assessment that involves all organization, and all its activities, instead of the simply emphasis in information security.

In this context is not important what standard is selected to apply, because both can be used to support security requirements and both can be applied to all type of organizations, no matter what dimension they have.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Definition of Information Security Policies

According to Merkow and Breithaupt (2014), information security measures try to address at least one of three goals- that are considered the basis of all security programs. These are: "Protect the confidentially of data; Preserve the integrity of data; and Promote the availability of data for authorized use" (p.20).

Information security can generally be understood as the set of measures that constitute security policy, by which the organization must determine which assets may be materially affected, i.e. define what the value of an asset is.

In this context, the policy is a "set of organizational guidelines that dictates certain behavior within organizations" (Whitman & Mattord, 2010, p.14), a situation that puts the focus of the information security in people and their attitudes when dealing with information. On the other side, Smith (2013) refer to information policies as security requirements, i.e., as the description of the protection that organizations need.

The information security policy, in accordance with the codes of good practice, is supported by a document that should be known by all employees of the organization (International Organization for Standardization, 2013b). This document shall contain:

- The commitment of management;
- The definition of what is meant by information security;
- The description of policies;
- The definition of responsibilities for information security;
- References to support documentation.

Another perspective on security policy refers to it as a high-level plan that outlines the objectives of these procedures, defining the details of implementation policies, i.e. providing a description of security in general (Barman, 2002).

In this sense, assuming a policy as a high-level declaration that should be as brief and concise as possible, there must be no doubt about the intention of this policy.

To meet the security requirements specified for an organization, a good policy must:

- Be easy to understand;
- Be applicable;
- Be amenable to implementation;
- Be proactive;
- Avoid absolute terms;
- Meet business objectives.

On the other hand, it is important to consider the aspects linked to security checks, considering the legal implications and best practice to use.

Relatively to controls with legislative implications, can be identified the following (International Organization for Standardization 2013b):

- Intellectual property rights;
- Safeguarding of organizational records;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Protection and privacy of data of personal information.

In regard to the control of good practice, the following deserve special relevance (International Organization for Standardization 2013b):

- Information security policy document;
- Allocation of responsibilities in information security risks;
- Training and education in information security;
- Reporting of security incidents;
- Business continuity management.

In this context, before security policies being written, it is necessary to determine the general objectives, namely what business activities should be protected, or whether such protection should serve only the data flows or the information systems, where the first step is to determine what should be protected and why (Barman, 2002).

The information security policy is supported by a document that should be known by all employees of the organization, and it specifies the management commitment, what organization understands as information security, what are the policies, who are responsible for information security, and also related documentation, as referred by International Organization for Standardization (2013b).

Solutions and Recommendations

The understanding of the relevance of information in the development and performance of organizations depends, in essence, of the understandings that the key decision-makers have about it and of their articulation with the organizational culture.

If we consider that the organization is competing in a global environment where information is acquiring greater relevance, the decision about choosing among the panoply of information at the disposal of organizations can make a difference in terms of performance, because the value creation depends more on the use of information than on the information systems and ICT that support it.

The economic and social consequences resulting from unemployment and low training on the use of nowadays tools (mainly based in information systems and ICT), may be proved catastrophic for the development of organizations and society, many times contributing to widen the gap between the more developed countries with stronger economies.

This results in a reduced ability to withstand competition from global enterprises and major inequalities that are brought by the emergence of a society that is based on the information.

The permanent learning depends in large degree of all involved and interested in their own processes and their adequacy to the culture and organizational goals.

The connection of people to organizations relates closely with the ability that they have to meet the needs and desires of human capital. There are thus not only implied compensation issues, a factor often seen as the only one relevant to limit individual performances, but also social situations and self-esteem, seeking the development of people in all its aspects.

The organizational culture and the organization's ability to respond to the people aspirations constitute the foundations of connection that is needed between people and organizations, with the search for balance between needs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Human resources or people have only to do with the framing in various areas of analysis. Human resource if the scope is restricted to working relationships, people if is considered a broader social and professional environment.

In a matter of security, particularly in information security, people are the final component but are indeed the most important one. This statement assumes a wider relevance, because people are present in every activity related with healthcare; as end-user of information systems and ICT, as patients, as medical or administrative staff, or even as suppliers of equipment or medicines, essentially because all of them deal with information.

In this sense, Wallin and Xu (2008) states that "information is a valuable asset in the information society, particularly in the healthcare sector where access to patient information is one of the most important factors for safety and quality" (p.13), referring that patient information, being sensitive, must be classified as confidential.

It is also important to take care about those that deal with patient's information during their normal activities, essentially because people deal with huge amount of information that must be confidential, integer and available.

The sensitivity of information that remains available, by information systems and ICT usage, and is frequently accessible by people in organizations, implies a proper approach to the measures that must be taken to protect it. To control accesses and handling only by authorized user when needed, it´s important, according Gaiveo (2011), to understand the relations that occur between information, people and security.

When the amount of personal health information is growing in all segments of healthcare (Wallin & Xu, 2008), it is important to assume that the need for support the creation, store and exchange of that information is an organizational imperative. By that, one of the main problems with healthcare become information systems that are currently used, is the storage of information in different formats, because some are open and other proprietary.

This diversity of formats is a serious impediment to sharing of information between all those involved, and may even become a critical factor in the protection of shared information due to vulnerabilities that originate the necessary integration between different systems.

Being "the main purpose of information security in healthcare is mainly to achieve two important aims: high-leveled patient safety and patient privacy" (Wallin & Xu, 2008, p.14), it's important to achieve a high level of patient safety in order to provide him/her with the best care as possible, using in the appropriate ways their information.

The needs of organizations in the establishment of an information security policy can be mitigated by the use of international standards for safety of information previously submitted, whereas a set of complementary activities is:

- **Identify and Classify Assets:** Achieved by using Annex B (Identification and valuation of assets and impact assessment) of ISO/IEC 27005 (International Organization for Standardization, 20011), that specify assets in primary and supporting. Primary assets are: Business processes & activities, and Information, while supporting assets are: Hardware (desktops, tablets, etc.), Software (applications, system, etc.), Network (switches, bridges, etc.), Personnel (end users, developers, helpdesk staff, etc.), Site, and Organization's structure;
- **Identify and Typify Threats:** Provided by Annex A (Threats to health information security) of ISO 27799 (International Organization for Standardization, 2008), but shouldn't be restricted to it, being a good option take into account threats source, in a similar way that is done in ISO/IEC 27005

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in their Annex C (Examples of typical threats), that presents threats by type, namely: Physical damage, Natural events, Loss of essential services, Disturbance due to radiation, Compromise of information, Technical failures, Unauthorised actions, and Compromise of functions;

- **Identify Vulnerabilities that Could Be Explored by Threats:** Done by applying Annex D (Vulnerabilities and methods for vulnerability assessment) of ISO/IEC 27005 (International Organization for Standardization, 2011), that presents vulnerabilities and threats that could explore each one. The examples of vulnerabilities versus threats are presented by type as: Hardware, Software, Network, Personnel, Site, and Organization;
- **Establish a Risk Assessment:** Through the Annex E (Information security risk assessment approaches) of ISO/IEC 27005 (International Organization for Standardization, 2011). With the aim of make the risk assessment according to the likelihood of occurrence of threat and ease of exploitation, taking into account the value of assets and business impact, in order to rank the threats by measuring the risks;
- **Planning and Implementing a Security Policy:** What should be done by applying ISO 27799 (International Organization for Standardization, 2008) guidelines, the same that are provided by ISO/IEC 27002, but within the specification of healthcare information security issues.

Besides that, controls presented by ISO 27799 contain modifications related with healthcare comparatively to ISO/IEC 27002 controls. Being necessary that, in addition to following the guidance given by ISO/IEC 27002, health organizations should follow the specific controls and implementation guidance provided by ISO 27799.

From that, emerges that those standards, and the appropriated application of their annexes, constitutes a fundamental resource that, jointly with ISO/IEC 27001, and also with ISO/IEC 27002, are fundamental to any information security program that intends to implement an ISMS.

CONCLUSION

In fact, no matter what experience people have about security or even what kind of security policies are in place; what seems to be undeniable is that information and its support require new attitudes and approaches to deal with the new contexts that are on the top of common preoccupations.

This need to ensure the security of business information is something that often is not the target of attention of organizational leaders. When the information it is assumed as an important organizational resource, business information can make the difference in the acquisition of competitive advantages and boosting the development of organizations.

All this context makes obvious the role of individuals in the performance of organizations, particularly when of those depends on the proper use of resources to respond appropriately to new business requirements. Before the biggest competitive demands that organizations face, where the information seems to emerge as a feature of considerable value and indispensable in the various organizational environments, people provide the skills and knowledge appropriate to the acquisition of competitive advantages for organizations.

The attraction of attention to the various organizational requirements in terms of business activities, is only achievable if the attention is given to the various kinds of information that the organizations utilize in their day-to-day routine; while ensuring their confidentiality, integrity, and availability according to the specific needs of the organization.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Moreover, supporting information flows within the organization and between this organization and its business partners; the information systems and ICT should be the target of similar attention of organizational leaders, including the safety aspects considered in terms of policies definition and their implementation in the organization.

Note that security concerns must be considered in a holistic manner, in order to incorporate into the organizational culture, looking for so that all actors in the organizational activities are sensitized about their relevance to the activities developed by the organization.

In this context, the creation of a climate of confidence is crucial to the pursuit of such activities. Essentially on the assumption of a responsible environment involving, as stakeholders, all participants in organizational activities, regardless of their placement are internal or external to the organization.

Any solution adopted by organizations in these areas will involve people as active parties, ensuring that the information security policies correspond to the needs of different stakeholders respecting the organizational culture and the applicable legislation. This strand is indeed a fundamental angle in the success of security measures to be adopted by organizations, seeking thus to avoid overrunning of the controls that are implemented according to specified requirements.

Together with this, organizations should apply the recommendations contained in the information security standards such as ISO/IEC 27002, to be used simultaneously with the other standards of series ISO/IEC 27000, that should be applied in conjunction with specific standards such as ISO 27799 in the context of information security in the healthcare sector.

These recommendations, jointly with the identification of potential threats that can take advantage of vulnerabilities in terms of information security, should allow the implementation of security policies, but also the raising of people awareness about the problem of information security, ensuring the protection of information in an appropriate manner to the needs of people and organizations.

REFERENCES

Barman, S. (2002). *Writing information security policies*. Owensville, MO: New Riders Publishing.

Bowen, P., Hash, J., & Wilson, M. (2006). *Information security handbook: A guide for managers.* NIST, Gaithersburg, MD. doi:10.6028/NIST.SP.800-100

Gaiveo, J. (2011). SMEs e-Business security issues. In M.M. Cruz-Cunha & J.E. Varajão (Eds.), Innovations in SMEs and conducting e-Business: Technologies, trends and solutions (pp. 317-337). Hershey, PA, USA: IGI Global. doi:10.4018/978-1-60960-765-4.ch018

HIPAA, Health Insurance Portability and Accountability Act 1996. Public Law 104-191. (1996). Senate and House of Representatives of the United States of America.

International Organization for Standardization. (2008). *ISO 27799 – Health informatics - Information security management in health using ISO/IEC 27002*. Geneve, Switzerland: ISO.

International Organization for Standardization. (2009). *ISO 31000 – Risk management – principles and guidelines*. Geneve, Switzerland: ISO.

International Organization for Standardization. (2011). *ISO/IEC 27005 – Information technology - security techniques - information security risk management*. Geneve, Switzerland: ISO/IEC.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

International Organization for Standardization. (2013). *ISO/IEC 27001 – Information technology - security techniques - information security management systems – requirements*. Geneve, Switzerland: ISO/IEC.

International Organization for Standardization. (2013). *ISO/IEC 27002 – Information technology - code of practice for information security controls*. Geneve, Switzerland: ISO/IEC.

International Organization for Standardization. (2014). *ISO/IEC 27000 – Information technology - security techniques - information security management systems – overview and vocabulary*. Geneve, Switzerland: ISO/IEC.

Merkow, M. S., & Breithaupt, J. (2014). *Information Security: Principles and Practices* (2nd ed.). Indianapolis, IN: Pearson IT Certification.

Peltier, T. R. (2010). *Information security risk analysis* (3rd ed.). Boca Raton, FL: Auerbach Publications. doi:10.1201/EBK1439839560

Scholl, M., Stine, K., Hash, J., Bowen, P., Johnson, A., Smith, C. D., & Steinberg, D. I. (2008). *An introductory resource guide for implementing the health insurance portability and accountability act (HIPAA) security rule (Rev. 1)*. NIST, Gaithersburg, MD. doi:10.6028/NIST.SP.800-66r1

Smith, R. E. (2013). *Elementary Information Security*. Burlington, MA: Jones & Bartlett Learning.

Stoneburner, G., Goguen, A., & Feringa, A. (2002). *Risk management guide for information technology systems.* NIST, Gaithersburg, MD.

Wallin, E., & Xu, Y. (2008). *Managing information security in healthcare: A case study in region Skåne* [Unpublished Master Dissertation]. Lund University, Lund, Sweden.

Weaver, R., Weaver, D., & Farwood, D. (2013). *Guide to network defense and countermeasures* (3rd ed.). Boston, MA: Thomson Course Technology.

Whitman, M. E., & Mattord, H. J. (2010). *Management of information security* (3rd ed.). Boston, MA: Thomson Course Technology.

Whitman, M. E., & Mattord, H. J. (2011). *Principles of information security* (4th ed.). Boston, MA: Thomson Course Technology.

Workman, M., Phelps, D. C., & Gathegi, J. N. (2013). *Information Security for Managers*. Burlington, MA: Jones & Bartlett Learning.

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 5(4); edited by Anastasius Moumtzoglou; pages 15-33; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 85
The Politics of Establishing ICT Governance for Large-Scale Healthcare Information Infrastructures

Gro-Hilde Ulriksen
University Hospital North Norway, Norway

Rune Pedersen
University Hospital North Norway, Norway

Gunnar Ellingsen
Arctic University of Norway, Norway

ABSTRACT

In Norway, the focus on interoperability and communication across healthcare practices has increased the need to connect ICT portfolios at different levels of healthcare, into large-scale information infrastructures (II). Governing healthcare practices is exceptionally complex, due to the diverging goals and policies of the heterogeneous actors involved. Establish well-functioning ICT governance organizations to handle these large infrastructures is therefore important. Using information infrastructure theory, and governance literature from the IS field, this paper contributes with empirical insight to the longitudinal and political process of establishing ICT governance in a healthcare context, reporting from one of Norway's largest health ICT projects, situated in the North Norway Regional Health Authority in 2012–2016. Our focus was on the following research questions: How does organizational politics shape the process of establishing an ICT governance organization in a heterogeneous healthcare environment, and what does it take to establish such ICT governance organization?

DOI: 10.4018/978-1-7998-1204-3.ch085

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Standardization of technology and work processes, to reach seamless integrations and semantic interoperability in Norwegian healthcare, has gained increased focus over the last years. The growing need for inter-organizational collaboration (Croteau, Bergeron, 2009; Dahlberg & Helin, 2014) and communication has raised the need for regional information and communication technology (ICT) portfolio. The role of the EPR systems, moving from local information storage systems, to large-scale user-centered work tools, has been particularly important. Consequently, the ICT portfolios have expanded in size and complexity. Hence, well-functioning ICT governance organizations at different levels of healthcare practices has gained increased focus. ICT governance include how to design and implement effective organizations by creating flexible ICT and information system (IS) structures and processes (Patel, 2002). The overall goal is for governance organizations to ensure successful deliverance of healthcare services (Beratarbide & Kelsey, 2009). There are increasing evidence related to establishing a connection between well-organized governance of health care organizations, and improved organizational performance (Tabish, 2012). However, hospitals and health systems struggle with matters of governance, particularly related to care standardization, and quality improvement (Tabish, 2012).

When making an effort to regionalize and standardize ICT portfolios, ICT governance organizations are important for maintaining the regional focus and handling challenges along the way. Traditionally ICT governance organizations were run in a top-down manner (Weill & Ross, 2004), this has however rarely proven efficient or successful for heterogeneous healthcare practices (Constantinides & Barrett, 2014; McGinnis, 1999). Therefore, a pressing need for shifting to a more bottom-up governance structure, focusing on the dynamic interactions between technical and social elements in ICT design (Constantinides & Barrett, 2014) has raised. Given the increased ambitions of information sharing, healthcare is characterized as institutions with different goals and policies, different ICT portfolios in play, and, stakeholders with different interests. Hence, it is necessary to look at the complexity of ICT governance, and the challenges of governing ICT portfolios at regional levels of healthcare. Introducing inter-organizational governance is an attempt to overcome the lack of interoperability and standards in healthcare (Dahlberg & Helin, 2014).

The contribution of this paper is to provide empirical insight to the longitudinal and political process of establishing an ICT governance organization within a healthcare context. Based on this, we ask the following research questions: How does organizational politics shape the process of establishing an ICT governance organization in a heterogeneous healthcare environment, and what does it take to establish such ICT governance organization?

We have gathered our empirical data, by following the steps of a regional initiative in the North Norwegian Health Authority. In 2012, this health region completed a large tender, and decided to regionalize their new ICT portfolio. To carry out these changes they established a regional project (dubbed BigProject), to run from 2012-2016. BigProject was one of the largest ICT investments in Norwegian healthcare, and the main goal of the project was to establish a regional ICT portfolio as a foundation for regionally standardized patient pathways, decision support, and integrations between clinical ICT systems (Christensen & Ellingsen, 2013). A regionalization, including standardizing EPR workpractice, was necessary requirements for reaching such goals, and for enabling the Health Authorities to better administrate and compare the hospitals in the region. In addition, the BigProject worked in close collaboration with the largest EPR vendor in Norway, on developing a more structured and interoperable EPR system, in order to communicate across heterogeneous healthcare practices (Nasjonal-IKT, 2012).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The data was collected by using a qualitative interpretive method (Klein & Myers, 1999; Walsham, 1995), including open-ended interviews, document studies, and participation in meetings and workshops. Through this approach, we aimed to emphasize various viewpoints of the process in order to achieve a deeper understanding of the challenges detected.

Theoretically, we applied information infrastructure theory; see (G. C. Bowker & Star, 2000; Hanseth & Lyytinen, 2010; Hanseth & Monteiro, 1998; Hanseth, Monteiro, & Hatling, 1996; Star & Ruhleder, 1996) frequently used to characterize and analyze large-scale integrated information systems portfolios (Garrod, 1998; Meum, Monteiro, & Ellingsen, 2011), and the interconnection between users and technology in heterogeneous healthcare practices (Hanseth & Lyytinen, 2010; Hanseth & Monteiro, 1998). We also used ICT governance literature from the information systems field; see (Beratarbide & Kelsey, 2009; Brown, 1997; De Haes & Van Grembergen, 2005; Simonsson & Johnson, 2005).

The rest of the paper organized as follows. First, the theoretical framework is introduced. Next, we present and elaborate on the method. Further, the case, including the BigProject and the new EPR is describe. We then present a discussion, emphasizing on different governance perspectives, and methods of structuring an ICT governance organization, aimed at managing large-scale information infrastructures. Last, we conclude the paper.

THEORY

The goals of integrated care, evidence-based treatment and standardized patient pathways have led healthcare organizations to invest heavily in integrated ICT systems (Chantler, Clarke, & Granger, 2006; Chiasson, Reddy, Kaplan, & Davidson, 2007; LeRouge, Mantzana, & Vance Wilson, 2007). Accordingly, we are not dealing with just one system, but instead with a portfolio of interconnected systems across institutional, departmental, and professional boundaries. As a way of conceptualizing this idea, the notion of information infrastructure (II) is particularly useful (G. Bowker, Timmermans, & Star, 1996; Ellingsen, Monteiro, & Munkvold, 2007; Star & Ruhleder, 1996; Timmermans & Berg, 2003). From a technical viewpoint, assembling an II involves designing, implementing, integrating, and controlling increasingly heterogeneous ICT capabilities (Tabish, 2012). Socially, creating an II requires organizing, and connecting heterogeneous actors with diverging interests, in ways that allow II to grow and evolve. Furthermore, IIs are heterogeneous, and open to an unlimited number of participants, such as users, vendors, and technical components (Hanseth & Lyytinen, 2010; Hanseth & Monteiro, 1998), which is important for use in a healthcare setting. A particularly important concept is the installed base (Hanseth & Monteiro, 1998), which implies that an II never develops from scratch, but emerges and evolves from an existing installed base. When developing new EPR systems, considering the old portfolios and practices in Health Trusts is important. For the different parts of an II to communicate, standards are core elements (Hanseth & Lyytinen, 2010). Standards ensure high-quality care through best practices development (Timmermans & Berg, 2003), increased efficiency, as well as ensuring seamless patient trajectories over organizational borders (Pedersen, Meum, & Ellingsen, 2012).

A pressing question is how and to what degree an II in general and standards in particular can be managed at different levels of healthcare. In the II literature, several authors; see (Edwards, Bowker, Jackson, & Williams, 2009; Karasti, Baker, & Millerand, 2010; Pipek & Wulf, 2009), have used the notion of infrastructuring, in order to emphasize the proactive engagement with large ICT portfolios. These insights are relevant for understanding the mechanisms for change; however, there has been less

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

focus on the more formal governance of organizational structures and configurations of II's. There is an increased need to establish ICT governance organizations that make decisions, as well as monitor results, and performances (Beratarbide & Kelsey, 2009), at different healthcare levels. Our focus in this paper is on interorganizational ICT governance at a regional level.

Most information systems had in-house ICT governance until the mid-1990s. Therefore, ICT governance has often been applied from an internal perspective (Boynton, Jacobs, & Zmud, 1992; Brown, 1997; Brown & Magill, 1994; Sambamurthy & Zmud, 1999). It is thus challenging to establish regional interorganizational ICT governance. ICT governance specifies the decision rights and accountability framework to encourage desirable behavior in ICT usage (Weill & Ross, 2005). We adhere to the following definition of ICT governance: "The preparation for, making, and implementation of decisions regarding goals, processes, people, and technology, on a tactical and strategic level of the IT organization" (Simonsson & Johnson, 2005). Stars and Ruhleder (1996) stated that the configuration mechanisms of governance are typically a mixture of various structures, processes, and relational aspects (Star & Ruhleder, 1996). Implementing ICT governance contributes to ensure successful delivery of healthcare according to Beratarbide & Kelsey (2009). The overall goal of an ICT governance organization is "to assure the stakeholders that things will go as expected, and ensure the successful delivery of healthcare services" (Beratarbide & Kelsey, 2009). Many ICT-related management frameworks, methodologies, and standards are used today; see (Beratarbide & Kelsey, 2009; Van Grembergen, De Haes, & Guldentops, 2004). None forms a complete ICT governance framework, but all have a role in assisting organizations towards more effectively managing and governing their information and related technologies (Beratarbide & Kelsey, 2009), as well as identifying ICT governance weaknesses.

Previously, a top-down approach with a clear ICT governance structure defining necessary decisions, and who should make them, was frequently used (Weill & Ross, 2004). Management studies promoted this design based on pre-defined models of work practices see, for instance, (Ashkenas, Ulrich, Jick, & Steve, 2002; Davenport, 1993). However, such strong controlling ICT governance for defining and making decisions (Weill & Ross, 2004), has been ineffective, and even impossible to apply to II in healthcare (Constantinides & Barrett, 2014). Several actors on the clinical and technical sides of healthcare need to be included in such governance. Due to the constant growth in complexity and deviation from original intentions, an II is impossible to govern completely in a top-down fashion (Croteau & Bergeron, 2009; Hanseth & Lyytinen, 2010).

In interorganizational contexts (such as in our case), Van Grembergen et al. (2004) suggest that ICT governance should include cooperation mechanisms to improve coordination of stakeholders with different ICT backgrounds (management and governance histories), and competence (ICT assets and resources) (Dahlberg & Helin, 2014). This because interorganizational relationships mature dynamically, and collaboratively, over various states (Croteau & Bergeron, 2009). The purpose of such governance is to ensure that organizations like Health Trusts have structures, processes, and mechanisms for collaboration, resolving disagreements, and organizing work on the interorganizational and organizational levels (Dahlberg & Helin, 2014). Improved quality and more interoperable health information is necessary, but very challenging to match with ICT governance principles and benefits in large scale interorganizational IIs.

However, despite establishing cooperation mechanisms, the size and scope of an II may be a serious challenge to ICT governance. Heterogeneous stakeholders have different goals and strategies for reaching them, resulting in frequent tension. This is particularly evident in a healthcare context. As a result, regionalization processes may be extremely challenging to accomplish. An obvious challenge is the tension between standardization and flexibility, recognized by Hanseth et al. (1996). However, flexibility

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

is necessary at the local level of a healthcare II, which enable users to work efficiently. In contrast, the regional perspective emphasizes a need for standardization, and the ability to compare different units as a part of running a more efficient and cost-effective healthcare services. In this regard, Constantinides and Barrett (2014) suggest a polycentric governance approach in which different stakeholders are engaged in dynamic and adaptive governance processes (Constantinides & Barrett, 2014).

Polycentric governance includes organizing a number of governing units at diverging levels, instead of one monocentric governance unit (McGinnis, 1999). In such governance model, there is a distribution of decision-making across organizational layers, and among a broad range of stakeholders, where each layer deals with associated subjects at a gradually larger scale and less-detailed level (McGinnis, 1999). This way, different actors in an II (such as a Health Trust) participate in ICT governance by controlling parts of an ICT portfolio. Consequently, this may lead to a smoother regionalization process in which the actors do not need to give up all local control. One key advantage of polycentric governance is the possibility of creating general formed rules that can later be adapted to specific local needs (McGinnis, 1999).

However, this model does not represent a "fast track to salvation;" it requires that actors spend extensive time and energy on negotiating and compromising on acceptable collaborative solutions (Latour, 2005). Even in cases where a common forum is established, it may be impossible to agree on governance structures acceptable to all parties, because of the heterogeneity of interests and resources involved in healthcare IIs (West, 2007). In order to grasp the challenges of governing information infrastructures, it is crucial to understand the various interests and associated mechanisms and how they play out over time (West, 2007).

METHOD

This study is positioned within a qualitative interpretive paradigm (Klein & Myers, 1999). It contributes to a longitudinal qualitative study connected to a large-scale EPR project in the Northern Norway Regional Health Authority. Analysis of longitudinal research is a continuous and iterative process, with an ever-changing intensity, focusing on developing and increasing the understanding of a phenomenon, by exploring diverse viewpoints within a specific context (Klein & Myers, 1999; Walsham, 1995).

This North Norwegian Health region consists of five Health Trusts including 11 hospitals. The regional EPR project (BigProject) has 25 employees and 5 sub-projects, of which two focuses on ERP. We interviewed participants from Health Trusts, local and regional management organizations, Regional Health Authorities, and BigProject to establish different viewpoints of the regional processes. Since regional standards had not yet been implemented in the Health Trusts, interviewing physicians and other end-users was not required at this point. Their perspectives will be more important to enlighten after completing the implementations of standards in 2015–2016.

The fieldwork builds on the first author's role, working in the standardization of practice project for two years, in addition to participating in workshops, discussions, and meetings in this project, and also the development of the new EPR. The data collection includes 11 open-ended interviews conducted in 2014–2015. Except from one, all the actors asked agreed to be interviewed. The interviews lasted 60–120 minutes each. The interview guide was semi-structured with questions related to the standardization of practice projects. This included pros and cons of standardization, why regionalization is important, and what regional challenges remain after the projects are finished. There were also questions concerning challenges with establishing and organizing a regional ICT governance organization.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Overview of interviewed participants

Informant	Time	Background
Four informants from BigProject	60–90 min	Laboratory, nursing, health secretary, technical
Three informants from regional ICT management	60–120 min	Technical, nursing.
Two local project leaders from Health Trusts	60 min	Nursing, occupational therapy
Leader from the local governance organization, UNN	60 min	Nursing
Leader from a Regional Health Authority	90 min	Economy

Refer to Table 1. The first author transcribed the interviews, and analyzed, and categorized the text into main issues in collaboration with the other authors. We used the hermeneutic circle, moving from understanding parts of the process to understanding the whole process (Klein & Myers, 1999). The challenges of organizing a regional governance organization, and how to deal with tension between different actors, were some main issues. Then we discussed and reflected on these subjects in relation to the context, the theory, and the research questions. In addition, we supplemented the data analysis with project documents and reports to acquire the best possible outline of the processes.

CASE

The Regional Strategy

After a prolonged bid for tender process in 2011, the North Norwegian Health Authority decided to invest in new clinical ICT systems for all their 11 hospitals, and at the same time regionalize the ICT portfolio. As a result, the Health Region established BigProject for the 2012–2016 period. With a cost likely to exceed €100 million (Christensen & Ellingsen, 2013), this project constitutes one of the largest and most ambitious ICT projects in Norwegian healthcare. The goals of the project were to establish

- Standardized clinical pathways
- "Best practice" standards for EPR workpractice and procedures across the Health Region
- Clinical decision support
- Regional integrations and interoperability between clinical systems

A crucial part of the project was contributing to developing a new EPR system (NewArena), using a completely new architecture with a user-centered configuration, and structured clinical content. As the hospitals in the Health Region already had an up-and-running EPR system (ClassicEPR), a crucial part of this project was to establish a strategy for a smooth transition between ClassicEPR and NewArena. In the next sections, we elaborate in more detail on these efforts.

In this process, BigProject had the primary responsibility for handling the major activities. However, the existing regional ICT management organization was also ascribed a crucial role. Their 250 employees were responsible for delivering ICT services, as well as running and maintaining the technical parts of the ICT portfolio in the Health Region. The regional ICT management collaborated closely with various

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

sized local ICT organizations in the different Health Trusts. In addition, both BigProject and the regional ICT management collaborated with a regional clinical ICT advisory board, established for handling the clinical content of the ICT portfolio at a regional level. The advisory board had representatives from all the five Health Trusts in the region, and met about once a month to provide strategic recommendations for the Regional Health Authorities. The advisory board acted as decision-makers for regional ICT issues, and thus represented a temporary de facto regional ICT governance organization. Even so, it was a common understanding in the Health Region to replace the advisory board with a permanent regional governance organization, which worked on a daily basis, when this organization was established.

Developing New EPR

In 2012, BigVendor started developing NewEPR based on openEHR architecture. OpenEHR built on standardized information models, open source components, and highly structured clinical content, including archetypes as core components (Beale & Heard, 2008). Archetypes are structured data elements of clinical concepts, envisioned to ensure technology-independent interoperability, easy reuse of information, and efficient decision support (Chen, Georgii-Hemming, & Åhlfeldt, 2009). Experienced clinicians are responsible for defining the archetypes, this way the clinical personnel control what content to structure, as well as how and when to do it. It is possible to build the entire internal structure of schemes, processes and decision support in the EPR, by organizing all data in archetypes, and combining them to form templates. There is a call for close regional collaboration, to maintain and build the clinical content, including patient pathways and archetypes for NewArena. The goal is to have a new EPR system working across institutional borders capable to follow the patients' entire trajectories (Pedersen et al., 2012).

To support this organizationally, the University Hospital North Norway (UNN) established a regional archetype governance organization in 2013 with 10 employees. This unit will support clinical users in creating and maintaining the archetypes at various levels (locally, regionally or nationally). The regional archetype governance will be organized at an overall regional level in the new governance organization, independent of clinical applications like EPR, radiology and laboratory.

The Health Authority needed to decide on a strategy to combine ClassicEPR with NewArena, since NewArena is developed and implemented in modules over at least five years while ClassicEPR is still operative. This made it necessary to be able to manage two different EPR systems simultaneously, as well as ensure that users could switch seamlessly between the systems. Therefore, the regional ICT management implemented a technical solution, to ensure that the old and new systems can interact seamlessly. "A context sync is installed, to make the switch between NewArena and Classic EPR seamless for system users. However, switching between two systems increases the risk of user errors, and technical complications" (representative from regional ICT management). This change is technically demanding, as well as challenging, for users, as attention to what platform they are working on is required to accomplish their work (Christensen & Ellingsen, 2014). After implementing more modules of NewArena, some functionalities will be in ClassicEPR and others in NewArena. In addition, both technology and use are very different between ClassicEPR and NewArena. An example described by Christensen and Ellingsen (2014) from the first pilot of NewArena illustrates this: The new system had structured data based on archetypes, but Classic EPR did not support structured text. Physicians created archetype-based documents in NewArena to plan operations. Then they sent these documents to ClassicEPR, and secretaries included the documents in the surgery-planning module. Structured data registered in NewArena then

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

became free-text documents, and the benefits of using archetypes disappeared completely (Christensen & Ellingsen, 2014). This indicated the risk of users not receiving any actual benefits of NewArena, until there are several modules implemented.

NewArena also calls for a close engagement related to user's practice, and more request for ICT support and governance closer to the end-users. This relates to the goal of designing clinical decision support and patient pathways, and the need to keep them constantly updated. However, this is an impossible task for the existing ICT management organization to provide, due to its regional role, and thus its distance from users. "The distance between the clinics and the regional governance organization is too far. It is cumbersome to get hold of the right person to solve a problem" (local project leader). To make this work, there is a need for a regional governance organization. In the first pilot of NewArena, the testing related to the surgery outpatient clinic confirmed that users would require close follow-up related to make the new functionalities useful for them.

The Standardization of Practice Project

To prepare organizationally for the new system, and attain the Regional Health Authority's goals of higher quality, efficiency and interoperability, regional standards and routines for EPR usage had to be established. Standards are important for a large II such as a Health Region to work optimally. Big Project's standardization of practice project was therefore initiated in 2012. The goal was to "increase quality and safety in patient treatment, through standardizing clinical practice related to EPR usage across the region" (Nilsen, 2013). More than 500 system users from all 11 hospitals in the region participated in mapping EPR usage and establishing regional standards in 2012–2013. The project identified existing work practices (the installed base) at all hospitals, and used a best practice principles to standardize work routines and procedures. In addition to system users, the EPR vendor, and both regional and local ICT management organizations participated in this process. The implementation of the regional standards will be completed in 2015–2016, including areas such as organizational configurations, journal structure, access control, laboratory work, and workflow (Nilsen, 2014). A consequence of the regional standards are extensive changes to the different hospitals' work practices. For example, after implementing the standards, a new form of access control will be required to use the EPR system. Today, there are great variations related to what users at the different hospitals can access automatically. For some, this new admission control will provide them with automatic access to less of the EPR than they have today. Another change is that all hospitals must register equivalent activities at the same organizational level, making information easier to compare at the regional level. Consequently, the organizational structuring of the hospitals has to be altered.

"Establishing the regional standards is a success, leading to better collaboration within the Health Region than ever before" (project leader, BigProject). However, for some standards it has been difficult to agree. One example relates to registration. The regional standard states: "Registration is connected to the patient's illness (Nilsen, 2014)." This means connecting all practice to a specialized field, such as ear, nose, and throat, or neurology. A project leader in BigProject states: "At a small hospital, activity registration can no longer be connected to the clinician doing the job, but rather the illness of the patient." Following the regional standard, all activities are registered at a regional level, since all hospitals in one Trust share the departments and sections connected to different specialized fields. This way, the Health Authorities improved their overview of activities related to each field. However, the hospitals themselves no longer have the same local overview. Local hospitals do not see the benefits of changing

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

their registration practice leaving them with less local control of their data. There have been several attempts trying to reach an agreement on this matter, without any luck so far. Similar challenges are likely to appear after implementing the standards, and the users' experience how their workdays are affected. Project leaders in the BigProject underline the need for a regional governance organization, to form a strategy for handling such issues, as well as others related to aberrations, modifying existing, or forming new standards. Without a governance organization, they fear that the regional structures will fall apart after BigProject is completed. "Regional decisions have already been effectuated, and there is a risk of them dissolving without the necessary regional governance in place" (project leader, BigProject).

Consequently, there is an increased need for a well-structured regional interorganizational ICT governance organization to follow up on the results of the standardization of practice project, and prepare for NewArena. Working out an overall goal for regional ICT management, including what to do and how to get there, is necessary. "It may take several years and the process may need to be done in steps, but to move forward, we need to have a strategy stating what the region wants this organization to encompass" (project leader, BigProject).

Establishing a New Regional Governance Organization

When the Health Region in 2011 decided to regionalize the ICT portfolio and established BigProject, they also started working on forming a regional ICT governance organization. However, it was difficult to negotiate the nature, form, and location of such organization, preferably having the authority to enforce standards on the various Health Trusts, as well as strategically manage a large ICT portfolio.

The experiences and lessons learned with the regional ICT management, was that this organization could not sufficiently support clinical practice, thus be responsible for regional ICT governance. There is a clear difference between ICT governance and ICT management. ICT management focuses mainly on suppling ICT services in addition to products, as well as managing ICT operations (De Haes & Van Grembergen, 2005). ICT governance is much broader, concentrating on performing along with transforming ICT to meet present besides future demands of the business and the customers.

Today, all Health Trusts, especially the largest one, have well-functioning local ICT management organizations, due to the need for user support and governance close to the workers. The local ICT management at the University Hospital also emerged as a de facto organization for clinical ICT management, preparing for the future clinical governance of NewArena, as well as being a key player preparing UNN for developing as well as piloting the new EPR system. "My guess is regional functional governance is placed in the biggest Health Trust, they already have an established organization for governing both NewArena and ClassicEPR" (leader, BigProject). Giving up local control and ICT management for a newly established regional ICT governance organization, did not seem like a tempting offer. However, even if the users agreed on the need for local clinical governance (i.e., the local ICT organization), the actual content of this organization was far from clear. At one point, this resulted in a territorial dispute between the local ICT governance at UNN, and regional ICT management. The regional ICT organization locked out the local ICT management from the administration tool in ClassicEPR (user management, access, role definitions etc.), with the argument that this was a technical issue. The local ICT management, however, claimed this tool was part of the clinical ICT governance, and demanded continued access. They also supported their claim by implicitly referring to the regional ICT management's managing role: "It is the Health Trust that own the system and decides who gets access to what" (leader, local ICT governance). This indicated a strong need for a regional ICT governance organization to handle similar issues in the future.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

After discussing several suggestions without reaching an agreement, UNN was in 2013 pointed out to complete a proposal for organizing regional ICT governance. In their proposal from 2014, they suggested a fragmented governance model in which each Health Trust were responsible for governing separate areas of the ICT portfolio on behalf of the others, meaning that one Health Trust would govern the EPR; one the laboratories, another radiology and so on. This model was established, based on the regional discussions: "Many discussions relates to establishing one small governance unit in each Health Trust, or one regional governance unit" (leader, local ICT governance organization). The solution presented by UNN distributed governance competence and local control across the Health Trusts. The Health Authorities piloted this solution for radiology in 2014–2015. After an evaluation, they will decide whether to apply this model to the rest of the regional ICT portfolio. This model is in many ways comparable to a polycentric governance model. Governance is spread out in the organization instead of assembled in one unit (McGinnis, 1999).

The informants were skeptic of applying such fragmented model to EPR governance, since the EPR is very complex, and has many integrations and interconnections with the rest of the ICT portfolio. It is necessary to include a large group of technical and clinical personnel to handle EPR governance, and a group working specifically with archetypes and NewArena. How to organize this and how to separate EPR related clinical content, is complicated to work out. "The systems are so closely interconnected, especially the EPR and the laboratory, that even if the Health Trust agrees to split their governance this seems impossible without compromising the interoperability" (project leader, BigProject). Even with all the skepticism, this model made it possible to define who was responsible for different clinical governance areas. "It seems like the best solution the region is able to agree on for the time being" (representative, Regional Health Authority). A fragmented governance model was at least a starting point for regional collaboration on governing an interorganizational ICT portfolio. In addition, establishing and updating regional patient pathways across Health Trusts, primary care and other institutional boundaries have to be extensive parts of the regional ICT governance.

DISCUSSION

De Haes and Van Grembergen (2005) states that ICT governance contains a mixture of various structures, processes, and relational mechanisms (De Haes & Van Grembergen, 2005). We agree, but we emphasize that governing an evolving II should be less concerned with creating uniform organizational structures for generating order, and focus more on a process for handling diverging political interests, managing tensions, as well as complex interdependences. This case has illustrated that establishing a uniform interorganizational governance regime is a formidable, if not impossible, task. Taking into account previous research on complex II (Bowker et al., 1996; Hanseth & Lyytinen, 2010; Hanseth & Monteiro, 1998; Star & Ruhleder, 1996), the current outcome is far from surprising. Modern IIs have become increasingly more interconnected and interdependent. At the same time, ICT systems, and in particular EPRs have been increasingly included in local practice. Thus, the infrastructure must serve regional as well as local interests simultaneously. In this case study, this is expressed by tailoring NewArena to the clinical practice through the new openEHR architecture. At the same time, the Regional Health Authority aims at standardizing practices across the Health Trusts in the region.

The two perspectives of the regional portfolio call for different sorts of governance: NewArena requires a governance regime grounded in local practice with high competence on how the new technology af-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

fects the clinician's daily work. Thus, consider how the first pilot of NewArena at the surgery outpatient clinic identified that users would need close follow-up engaging with the new functionalities. In contrast, Big Project's standardization of practice calls for a more authoritative governance regime. The regional standards and routines must be implemented in clinical practice; in addition, someone have to make sure that the users adhere to the standards for the standards to continue to evolve alongside clinical practice. At this point, the regional ICT advisory board and the regional ICT management organization are not equipped to handle such a governance task on a permanent basis, since they are management, not governance, organizations. There is a need for a more powerful interorganizational governance organization operating at an overall level on a daily basis, to support successful delivery of healthcare.

In addition, governing an II also implies taking into account and building on the existing ICT portfolio, the installed base (Hanseth & Monteiro, 1998; Nasjonal-IKT, 2012). The BigProject leaders have considered the existing system with their current strategy to plan for years of interplay between ClassicEPR and NewArena. This enables seamless integration of NewArena, by replacing ClassicEPR in a stepwise manner avoiding a "Big Bang" implementation. Operating the two EPR solutions simultaneously however, has created challenges, such as the example from Christensen and Ellingsen (2014) in which a process moved from NewArena back to Classic EPR resulting in an outcome that is the same as today; thus, no improvements in NewArena are detected for users. The benefits for users may not be noticeable until replacing most, or all parts of the old system.

From a governance perspective, a key point is that the interplay between the old and new portfolios require thorough technical insight. Simultaneously, such a technical perspective relates closely to clinical practice, since technical and clinical decisions affect each other. Consequently, governing an II implies understanding and including different interrelated areas, where each part requires a unique governance regime. The three governance perspectives introduced in this paper, the local, the regional, and the technical, are interconnected. Changing one area of the infrastructure may result in unpredicted changes to other areas (Vikkelsø, 2005). These interactions may cause tensions in addition to conflicts of interest. For example, pushing too heavily on a regional perspective and standardization, may have unforeseen effects on local flexibility. An example is how physicians at small hospitals lost local overview and authority of their patient data, when the standardization of practice project implemented regional standards. Consequently, things that are very useful from a regional perspective, becomes a liability for local practice. Accordingly, based on the preliminary results, there are indications that regional ambition has been pushed a step too far, as the different Health Trusts could not agree on how to organize a governance organization able to serve all the various needs. Given New Arena's heavy influence on future work practice (decision support, patient pathways, structured EPR content, etc.), it is not surprising that the Health Trusts sought highly competent local ICT governance close to clinical practice, instead of relating to a newly established regional ICT governance organization.

A strategy for solving the local and regional challenges has been to divide the governance into smaller domains such as EPR, radiology, laboratory, etc., where each domain is anchored in one of the Health Trusts instead of at the regional level. This solution is similar to a polycentric governance model, with multiple governing units, several overlapping arenas of authority and responsibility (Constantinides & Barrett, 2014). To some degree, this may solve the local/regional tension within the specific area in question. A key challenge with the fragmented governance structure suggested is defining the boundaries, then the areas of responsibility between the different units. Thus, several independent organizational units end up controlling areas of the same system, and the ICT management organization that delivers technical solutions have to collaborate with several Health Trusts. In addition, some parts of the ICT

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

portfolio, such as EPRs and laboratories, are so closely connected that separating them may seem impractical if not impossible. Another challenge may relate to two governance organizations disagreeing on a topic. Who decides the outcome? These borders are not clear-cut (as we have pointed out), and require cumbersome negotiation between areas of responsibility, thus making governance a complex task to handle. Such fragmented governance models have been impossible to agree on in the past, according to McGinnis (1999). Even in cases where a common forum was established, it has been impossible to agree on governance structures acceptable to all parties, since heterogeneous interests and resources are involved in a large scale II.

One of the goals of NewArena is to operate across institutional borders, to follow a patient's entire trajectory gathering all necessary patient information in one place. The fragmented governance solution suggested, made embracing NewArena extra complex. Governance of NewArena demands a regional focus on clinical content independent of system vendors, leading to an increased need for regional ICT governance. There is a call for close regional collaboration to maintain and model the clinical content, including archetypes and templates. It is necessary to take into account the goals of the new system, when the interorganizational ICT governance organization is established. This indicates that trying out the suggested polycentric governance model will potentially create room for endless negotiations plus a constant need for compromises. If choosing this model, a key success factor is to define clearly, who makes the final decisions when the different governance organizations disagree. If not, it will be damaging for the whole II, and it will be challenging to assure the stakeholders that things will go as expected. In addition, it is necessary to make an overall strategy securing the appropriate handling of all necessary areas of interorganizational governance.

CONCLUSION

Given the increased ambitions of information sharing, standardization, and interoperability, to ensure successful delivery of healthcare, it is crucially important to establish a well-functioning II. A key factor for success is establishing an ICT governance organization. We have discussed several challenges of how and to what degree an II in general and ICT portfolios in particular, can be managed, by focusing on organizational structures as well as processes on how to formally govern II's. We emphasized the longitudinal and political process of establishing an interorganizational ICT governance in a heterogeneous healthcare context.

The goal is for the EPR systems to evolve from today's information storage systems to interoperable user-centered work tools. Standardization of technology and work processes, as well as interorganizational governance collaboration is necessary to reach such goals. The regional ICT governance organization must operate on a daily basis; make decisions, as well as monitoring results and performance on different healthcare levels. In addition, standards must evolve alongside the EPR to be useful.

For such ICT governance organization to succeed the Health Authorities and the Health Trusts have to define an overall strategy for what they want to include in this organization. Interorganizational ICT governance has to include different aspects of governing the regional ICT portfolio, including local, regional, and technical aspects, plus the ongoing tensions among them. Regionalizing an ICT portfolio is challenging, even if representatives from Health Trusts, ICT management organizations and vendors participate in the process, the standards established might not always fit local needs. Therefore, a strategy

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

for handling regional disagreements, besides evaluating requirements for revising and adding standards, is also important to work out. In addition, it is necessary to define the structure of such ICT governance organization.

The fragmented governance model suggested for the North Norwegian Health Region is similar to a polycentric governance structure, which offers opportunities for organizing several governing units at diverging scales instead of one monocentric governance unit. This is an attempt to get a step closer to regional collaboration, and seems to be as far as the Health Trusts are ready to go at this point. Using this model, they keep some local control of their ICT portfolio. A fragmented governance solution demands close collaboration, and clearly defined borders between the different actors. This implies that a polycentric ICT governance structure may be an acceptable but challenging solution for a loosely connected II as a Health Region.

REFERENCES

Ashkenas, R., Ulrich, D., Jick, T., & Steve, K. (2002). *The boundary less organization. Breaking the chains of organizational structure*. San Francisco: Jossey-Bass. Wiley.

Beale, T., & Heard, S. (2008). *OpenEHR Architecture: Architecture Overview.* London.

Beratarbide, E., & Kelsey, T. (2009). eHealth Governance, a Key Factor for Better HealthcareE: Implementation of IT Governance to ensure better care through better eHealth. In *Ethical Issues and security monitoring trends in global healthcare: Technological advancements*.

Bowker, G., Timmermans, S., & Star, S. L. (1996). *Infrastructure and organizational transformation: Classifying nurses' work*. London: Chapman and Hall.

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge: The MIT press.

Boynton, A. C., Jacobs, G. C., & Zmud, R. W. (1992). Whose Responsibility Is IT Management? *Sloan Management Review*, *33*(4), 32–38. PMID:10120625

Brown, C. V. (1997). Examining the Emergence of Hybrid IS Governance Solutions: Evidence from a Single Case Site. *Information Systems Research*, *8*(1), 69–94. doi:10.1287/isre.8.1.69

Brown, C. V., & Magill, S. L. (1994). Alignment of the IS Functions with the Enterprise: Toward a Model of Antecedents. *Management Information Systems Quarterly*, *18*(4), 371–403. doi:10.2307/249521

Chantler, J. C., Clarke, T., & Granger, R. (2006). Information technology in the English National Health Service. *Journal of the American Medical Association*, *296*(18), 2255–2258. doi:10.1001/jama.296.18.2255 PMID:17090773

Chen, R., Georgii-Hemming, P., & Åhlfeldt, H. (2009). Representing a Chemotherapy Guideline Using openEHR and Rules. *Studies in Health Technology and Informatics*, *150*, 653–657. doi:10.3233./978-1-60750-044-5-653 PMID:19745392

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chiasson, M., Reddy, M., Kaplan, B., & Davidson, E. (2007). Expanding multi-disciplinary approaches to healthcare information technologies: What does information systems offer medical informatics. *International Journal of Medical Informatics*, *76*, 89–97. doi:10.1016/j.ijmedinf.2006.05.010 PMID:16769245

Christensen, B., & Ellingsen, G. (2013). Standardizing Clinical Pathways for Surgery Patients through ICT. *Paper presented at the European Workshop on Practical Aspects of Health Informatics (PAHI '13)*, Edinburgh, UK.

Christensen, B., & Ellingsen, G. (2014). Advances in Intelligent Systems and Computing: Towards a Structured Electronic Patient Record for supporting Clinical Decision Making. In Á. Rocha, A. M. Correia, F. B. Tan, & K. A. Stroetmann (Eds.), *New perspectives in Information Systems ans Technologies* (Vol. 2, pp. 297–306). Switzerland: Springer International publishing. doi:10.1007/978-3-319-05948-8_29

Constantinides, P., & Barrett, M. (2014). Information Infrastructure Development and Governance as Collective Action. *Information Systems Research*.

Croteau, A.-M., & Bergeron, F. (2009). *Interorganizational governance of information technology Paper presented at the Proceedings of the 42nd Hawaii International Conference on System Sciences*, Hawaii

Dahlberg, T., & Helin, A. (2014). Formation of Voluntary Inter-Organizational IT Governance for Healthcare and Social Welfare IT – Theoretical Background and Empirical Evaluation. *Paper presented at the Proceedings of the 8th International Conference on the Theory and Practice of Electronic Governance*, At Guimaraes, Portugal. Retrieved from http://www.researchgate.net/publication/274251924

Davenport, T. (1993). *Process innovation Reengineering work through information technology*. Boston: Harvard Business School Press.

De Haes, S., & Van Grembergen, W. (2005). IT Governance Structures, Processes and Relational Mechanisms: Achieving IT/Business Alignment in a Major Belgian Financial Group. *Paper presented at the Proceedings of the 38th Hawaii International Conference on System Sciences*, Hawaii. 10.1109/HICSS.2005.362

Edwards, P. N., Bowker, G., Jackson, S. J., & Williams, R. (2009). Introduction: An agenda for infrastructure studies. *Journal of the Association for Information Systems*, *10*(5), 364–374.

Ellingsen, G., Monteiro, E., & Munkvold, G. (2007). Standardization of work: Co-constructive practice. *The Information Society*, *23*(5), 309–326. doi:10.1080/01972240701572723

Garrod, S. (1998). How groups co-ordinate their consepts and terminology; implications for medical informatics. *Methods of Information in Medicine*, *37*(4-5), 471–476. PMID:9865045

Hanseth, O., & Lyytinen, K. (2010). Design theory for dynamic complexity in information infrastructures: The case of building Internet. *Journal of Information Technology*, *25*(1), 1–19. doi:10.1057/jit.2009.19

Hanseth, O., & Monteiro, E. (1998). *Understanding information infrastructure.* Retrieved from http://heim.ifi.uio.no/~oleha/Publications/bok.html

Hanseth, O., Monteiro, E., & Hatling, M. (1996). Developing Information Infrastructure: The tension between standardization and flexibility. *Science, Technology & Human Values*, *21*(4), 407–426. doi:10.1177/016224399602100402

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Karasti, H., Baker, K., & Millerand, F. (2010). Infrastructure time: Long-term matters in collaborative development. *Computer Supported Cooperative Work*, *19*(3-4), 377–415. doi:10.100710606-010-9113-z

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *Management Information Systems Quarterly*, *23*(1), 67–94. doi:10.2307/249410

Latour, B. (2005). *Reassembling the Social: An introduction to actor network theory*. New York: Oxford University Press.

LeRouge, C., Mantzana, V., & Vance Wilson, E. (2007). Healthcare information systems research, revelations and visions. *European Journal of Information Systems*, *16*(6), 669–671. doi:10.1057/palgrave.ejis.3000712

McGinnis, M. (1999). *Polycentric Governance and Development: Readings from the Workshop in Political Theory and Policy Analysis*. Ann Arbor, MI: University of Michigan Press. doi:10.3998/mpub.16052

Meum, T., Monteiro, E., & Ellingsen, G. (2011). The Pendulum of Standardization. *Paper presented at the ECSCW 2011 Proceedings of the 12th European Conference on Computer Supported Cooperative Work*.

Nasjonal-IKT. (2012). Tiltak 48: Klinisk dokumentasjon for oversikt og læring. [Action 48: Clinical documentation for overview and learning]. Retrieved from http://www.nasjonalikt.no/no/dokumenter/andre_publikasjoner

Nilsen, A. (2013). *Prosjektplan HOS Implementeringsprosjekt*. Retrieved from Nilsen, A. (2014). *Standardiseringshåndbok for bruk av DIPS i Helse Nord*.

Patel, N. V. (2002). Emergent Forms of IT Governance to Support Global E-Business Models. *Journal of Information Technology Theory and Application*, *4*(2), 33–48.

Pedersen, R., Meum, T., & Ellingsen, G. (2012). Nursing Terminologies as Evolving Large-Scale Information Infrastructures. *Scandinavian Journal of Information Systems*, *24*(1), 55–82.

Pipek, V., & Wulf, V. (2009). Infrastructuring: Toward an Integrated Perspective on the Design and Use of Information Technology. *Journal of the Association for Information Systems*, *10*(5), 447–473.

Sambamurthy, V., & Zmud, R. (1999). Arrangements for Information Technology Governance: A Theory of Multiple Contingencies. *Management Information Systems Quarterly*, *23*(2), 261–290. doi:10.2307/249754

Simonsson, M., & Johnson, P. (2005). Defining IT Governance – A Consolidation of Literature. KTH Royal Institute of Technology, Sweden. Retrieved from http://www.ics.kth.se/Publikationer/

Star, S. L., & Ruhleder, K. (1996). Steps toward and ecology of infrastructure: Design and access for large Information spaces. *Information Systems Research*, *7*(1), 111–134. doi:10.1287/isre.7.1.111

Tabish, S. A. (2012). Healthcare: From Good to Exceptional Governance. *JIMSA*, *25*(2), 147–149.

Timmermans, S., & Berg, M. (2003). *The Gold Standard: The challenge of evidence based medicine and standardization in health care*. Philadelphia: Temple University Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Van Grembergen, W., De Haes, S., & Guldentops, E. (2004). *Structures, Processes and Relational Mechanisms for IT Governance*. Retrieved from Hershey. doi:10.4018/978-1-59140-140-7.ch001

Vikkelsø, S. (2005). Subtle Redistribution of Work, Attention and Risks: Electronic Patient Records and Organisational Consequences. *Scandinavian Journal of Information Systems*, *17*(1), 3–30.

Walsham, G. (1995). Interpretive case studies in IS research: Nature and method. *European Journal of Information Systems*, *4*(2), 74–81. doi:10.1057/ejis.1995.9

Weill, P., & Ross, J. (2004). *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Boston: Harvard Business School Press.

Weill, P., & Ross, J. W. (2005). A Matrixed Approach to Designing IT Governance. *Sloan Management Review*, *46*(2), 26–34.

West, J. (2007). *The economic realities of open standards: Black, white, and many shades of gray*. Cambridge, UK: Cambridge University Pres.

This research was previously published in the International Journal of Social and Organizational Dynamics in IT (IJSODIT), 6(1); edited by Michael B. Knight; pages 48-61; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 86
Inhibitors of Physicians' Use of Mandatory Hospital Information Systems (HIS)

Heiko Gewald
Neu-Ulm University of Applied Sciences, Germany

Corinna Gewald
Neu-Ulm University of Applied Sciences, Germany

ABSTRACT

Physicians' use of information systems continues to be a highly relevant area of information systems research. Although numerous studies have investigated the relationship between hospital physicians and hospital information systems (HIS), a comprehensive framework for assessing the factors which inhibit the use of HIS has yet to be developed. To advance these efforts, this exploratory research takes a qualitative approach to investigate HIS use inhibitors based on interviews with 48 informants and observations from over 40 hours of workplace shadowing in two German hospitals. The findings show that focusing solely on the user does not lead to successful system deployment in hospitals. Rather, the road to success needs to take a combined approach focusing on the user, the process and the system. If these three building blocks are well aligned, correspond with the organizational context and are supported by strong leadership, medical personnel are much more satisfied with the HIS, which results in more effective use.

1. INTRODUCTION

Information technology (IT) is used widely in hospitals in developed economies all over the world. Although it seems to be logical to employ hospital information systems (HIS) and it is "…no matter of if, but of when…" (Wiggins, Peterson, & Moss, 2015) they will be employed, anecdotal evidence and discussions with practitioners often give the impression that HIS are not performing up to expectations: Proof that these systems provide a sustainable and measurable positive impact on patients and the healthcare industry is still outstanding.

DOI: 10.4018/978-1-7998-1204-3.ch086

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Numerous authors, such as Buntin, Burke, Hoaglin, and Blumenthal (2011) and Denis (2009), have discussed the expected positive benefits from HIS, including improvements in the quality of care, a decline in healthcare delivery costs, and the avoidance of unnecessary procedures. However, critical voices, such as Ash, Berg, and Coiera (2004) or Koppel (2016), also associate serious negative outcomes like unnecessary mortality and redundant expenditures with the use of HIS. Despite the lack of proof of benefits, hospitals in developed economies continue to invest large amounts of money in their IT systems. Actual figures are hard to get and vary depending on sources. According to a report by the European Commission, most European hospitals spent around 2.5 to 3% of their operating budget on IT services (Buddrus, 2013; European Commission, 2013). This investment is significant and not likely to decrease in the near future (IDC, 2015).

Considering the expected benefits of HIS and the investments into IT, one wonders why healthcare systems around the world have difficulties living up to expectations (Davino, 2011). This poses the question whether HIS is able to provide the promised benefits (Kellermann & Jones, 2013). The answer to this question is influenced by multiple factors and is thus highly complex. Inspired by personal experience and anecdotal evidence, this research aims to shed light on this question by assessing the role of the user. Medical personnel's acceptance of HIS has been identified as an important prerequisite for successful healthcare delivery (Tulu, Burkhard, & Horan, 2006; Walter & Lopez, 2008) and the anticipated benefits of a HIS will only materialize if users use the IT as designated (Engelbert & Graeml, 2015). As Abouzahra et al. put it: "…the benefits of HIS can only be reached if [the HIS] are used in practice." (Abouzahra, Guenter, & Tan, 2015, p.14).

Due to their central role in healthcare delivery, medical personnel is often considered a major obstacle to successful IS deployment in hospitals (Boonstra, Versluis, & Vos, 2014; Paré, Sicotte, & Jacques, 2006). As Lowenhaupt put it: "Physicians' adoption [of HIS] has long been considered 'the holy grail' of clinical information systems: critically important, but elusive…" (Lowenhaupt, 2004, p. 12). This leads to the guiding hypotheses of this study that physicians and HIS are not well matched. Buntin's 2011 literature review identifies initial quantitative support for this hypothesis.

This paper delves deeper into this hypothesis, addressing the research question: What are the key factors that inhibit physicians from having a positive attitude towards hospital information systems in a mandatory use setting?

HIS are mandatory use systems, especially when used to document patient-related information. In a mandatory-use setting, whether a user likes or dislikes the IS does not determine whether or not he or she will use the IS. However, research and practice show that there are different levels of system interaction "…ranging from use to effective use…" (Burton-Jones & Grange, 2012, p.632). It seems obvious that a positive attitude towards the system is a vital building block to "more effective use" of the HIS. This research aims to provide further insights into why physician users of HIS do not always like working with the system and, more specifically, why physicians tend to not always use the systems as designed by system engineers for the mandatory-use hospital setting.

Numerous papers and several literature reviews (Abouzahra et al., 2015; Boonstra & Broekhuis, 2010; Boonstra et al., 2014; Buntin et al., 2011) identify a large body of research into the enablers but mainly inhibitors of hospital medical personnel HIS usage. Unfortunately, Boonstra et al. (2014, p.16) concludes that "…the literature is diffuse, and articles seldom build on earlier ones to increase the theoretical knowledge…" In light of this observation, this exploratory research investigates the root cause of the problem: the interaction between users and systems in the context of clinical work. Although numerous studies have already been conducted on the issue, there is still no comprehensive understanding and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

literally all papers call for further research into this issue. Instead of performing a further quantitative study, we build on available knowledge and conduct qualitative research to get at the root of the problem. Following a semi-structured interview guideline developed based on previously identified inhibitors of physicians' use of HIS, we held 48 interviews of physician users of HIS and shadowed such users for over 40 hours in two hospitals in Germany. The interviews and observations were coded, sorted and compiled to create a reference framework.

This paper is organized as follows: After reviewing the literature on IT adoption and usage in hospitals, we outline our research method. We then discuss our findings and derive the proposed framework. Finally, we discuss the limitations of our research, suggest further research and draw our conclusions.

2. LITERATURE REVIEW

Over the years, countless studies have tried to explain physician`s (non-)adoption of HIS. This section defines the relevant terms and discusses three major issues affecting the relationship between physicians and HIS in hospitals:

1. **Concepts of Usage:** The types of behavior physicians have displayed towards HIS
2. **Mandatory Use:** User behavior in environments where no choice can be made whether or not to use a system, i.e. where usage is not voluntary.
3. **Known Inhibitors:** Findings of previous research what inhibits users from using HIS.

2.1. Definitions

The term HIS is often interpreted differently. For this research, it is defined as the IT system(s) used by hospital physicians to document patient-related information. This is typically the electronic medical record; however, it can also be the general hospital information system (like the ERP for industrial firms) or another specialized system. The definition excludes function-specific medical IT systems such as x-ray machines, heart catheters etc. which are used for diagnostic purposes only (although the data may feed into other administrative systems). Our definition limits HIS to hospital-internal systems and excludes inter-organizational systems such as health information exchanges or electronic health records.

2.2. Concepts of Usage

2.2.1. Adoption, Acceptance, and Intention-to-Use

Numerous studies have investigated the interaction between users and systems in healthcare (see literature reviews by Boonstra and Broekhuis (2010), Holden and Karsh (2010), and Shaikh and Karjaluoto (2015)). The most popular objects of study are the electronic medical record (EMR) and HIS in general. Several popular IS adoption models have been utilized (such as TAM, TAM2, UTAUT and SCT), adapted to the healthcare context and sometimes extended to include specific constructs. Some researchers argue that specific traits of the physician's occupation (like the construct "perceived threat to professional autonomy" proposed by Walter and Lopez (2008). However, we have not found a study that sufficiently explains why physicians do not show the same adoption behavior as users in other industries.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

All the papers dealing with either adoption, acceptance, or intention-to-use (see Venkatesh and Davis (2000) for an in-depth discussion of these terms) share an inherent assumption that the use is, to some degree, free to choose whether or not to use the system. This is not the case for HIS in a hospital environment. Laws, policies and regulations dictate that healthcare providers must document their work extensively (Tulu et al., 2006). This means that if a hospital switches from a paper-based to electronic medical record keeping, physicians do not have a choice whether to use the system. They must use it, whether they like it or not.

2.2.2. Resistance

Scholars have observed that users often resist new IS implementations. This type of resistance is defined as "…opposition of a user to change associated with a new IS implementation…" (Kim & Kankanhalli, 2009, p.567). Users show active, passive, overt, or covert negative behavioral responses to avoid using a system (Laumer & Eckhardt, 2012). Several studies of resistance to IS implementation in general (Hirschheim & Newman, 1988; Kim & Kankanhalli, 2009; Lapointe & Rivard, 2005; Laumer & Eckhardt, 2012; Laumer, Maier, Eckhardt, & Weitzel, 2015) and of resistance to IS implementation in hospitals (Bhattacherjee & Hikmet, 2007; Doolin, 2004) highlight the importance of including the user in the implementation process, since changes to their work processes will impact them directly (Liu & Cheng, 2015). These studies find that resistance is a typical reaction to changes before or during the implementation of a new system. In the case of physician users of HIS, the vast majority of medical personnel in hospitals use systems that have already been implemented, so the concept of resistance is restricted to rather specific circumstances.

2.2.3. Continuous Use

To distinguish between first-time and ongoing use of an implemented system IS usage scholars have established the concept of "continuous use". Unfortunately, as pointed out by Gebauer et al. (2013), considerable research (e.g. Bhattacherjee, 2001; Karahanna, Straub, & Chervany, 1999) regards continuous use to be grounded in the same theoretical approaches used to explain initial IT adoption (i.e. TAM, UTAUT, TRA, TPB etc.). These theories use "intention to use" as a dependent variable, which is being increasingly questioned as the link between behavioral intention and actual usage is not as strong as often expected (Limayem, Hirt, & Cheung, 2007) especially in self-reporting research settings (Wu & Du, 2012).

Due to their focus on the intention construct, continuous use theories assume voluntary use of the system under observation. Logically, the few continuous use studies undertaken in the healthcare context focused on non-mandatory use systems (e.g. Tulu et al. (2006).

2.3. Mandatory Use

In summary, studies on adoption, resistance or continuous use do not help explain the behavior of hospital medical personnel using an existing system in compliance with hospital and statutory healthcare provision rules and regulations. It is mandatory that medical personnel use the HIS. In this case, the end-user has no choice whether or not to use the system (Melone, 1990). Management made that decision for all end-users by deciding to implement the system (Vehring, Riemer, & Stefan, 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Previous research shows that user behavior differs depending on whether system use is voluntary or mandatory (Gallivan, 2001). However, there is far less research into mandatory system use (Chan et al., 2010). Even though it is common in the business environment to require employees to use certain IS (Brown, Massey, Montoya-Weiss, & Burkman, 2002), research on the specific user behavior towards mandated use is scarce (Laumer, Maier, Eckhardt, & Weitzel, 2014).

Despite some discussion about whether system use can be really mandatory or whether a degree of voluntarism is always involved (DeLone & McLean, 2002; Vehring et al., 2011), scholars agree that the quality and intensity of use directly affects the degree to which system benefits are realized, even in required-use settings: "…even when use is required, variability in the quality and intensity of this use is likely to have a significant impact on the realization of the system benefits…" (DeLone & McLean, 2002, p. 5). This is due to the fact that even in mandatory settings the extent of system use varies by user (Burton-Jones & Grange, 2012; Melone, 1990).

Thus, theoretical models focusing on "intention to use" as a dependent variable are not as strong in mandatory context. While there will be variance in terms of use and workarounds, the mandatory-use setting per definition excludes user choice.

The question thus remains: What factors influence the degree of IS usage and what factors create positive or negative attitudes toward IS usage in a mandatory-use setting?

2.4. Inhibitors of HIS use

Table 1 below summarizes the factors identified by IS usage scholars which encourage or hinder physicians' use of HIS (Holden & Karsh, 2010). We added additional information in brackets to further clarify the original descriptions. Please note that Boonstra et al. (2014), Boonstra and Broekhuis (2010), Buntin et al. (2011), and Holden and Karsh (2010) are literature reviews. If these are stated as source they represent a secondary and not a primary reference.

3. RESEARCH METHOD

As discussed in more detail in the introduction, in order to get at the root question of what inhibits physicians from using an implemented HIS in a mandatory-use hospital setting, we conducted exploratory research using qualitative methods. Using identified inhibitors of physician's HIS use, we developed a semi-structured interview guideline and interviewed physicians in two hospitals in Germany and shadowed physicians to deepen our findings. The findings were coded, sorted and compiled to create a reference framework.

3.1. Research Objects

This study addresses the behavior of hospital medical personnel towards an information system provided by the hospital. We did not analyze the hospital as an organization which may benefit from HIS through improved cost-effectiveness or an improved doctor-patient relationship (geared towards a team approach instead of a 1:1 relation). Since organizational behavior is driven by different motives than individual behavior, we excluded this lens.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Inhibitors of physicians' HIS use

	Inhibitor	Source(s)
1	Lack of interoperability [the system is insufficiently connected with other HIS systems]	Kellermann and Jones (2013) Lærum, Ellingsen, and Faxvaag (2001)
2	Insufficient ease-of-use, system too complex [difficult to find information, insufficient GUI, etc.]	Kellermann and Jones (2013) Boonstra and Broekhuis (2010)
3	Insufficient efficiency within the system [too many clicks required, sign-on to multiple systems, etc.]	Boonstra et al. (2014)
4	Insufficient integration with other clinical processes [coding, billing, cross-departmental consultations, etc.]	Kellermann and Jones (2013)
5	Professionals need to adapt their working customs [change management issue]	Liu and Cheng (2015) Lærum et al. (2001) Shachak and Reis (2009) Wakefield et al. (2007)
6	Technology does not fit to professionals' needs / work procedures [system does not support the required functionality]	Villalba-Mora, Casas, Lupiañez-Villanueva, and Maghiros (2015) Heisy-Grove and Patel (2014) Boonstra et al. (2014) Tulu et al. (2006)
7	Threat to physicians' professional autonomy	Walter and Lopez (2008) Lowenhaupt (2004) Walter and Lopez (2008) Weeger and Gewald (2015) Chau and Hu (2002)
8	Negative impacts on physician-patient relationship	Shachak and Reis (2009)
9	Patient privacy and information security concerns	Thomas (2008) Gold, McLaughlin, Devers, Berenson, and Bovbjerg (2012) Villalba-Mora et al. (2015) Heisy-Grove and Patel (2014) Chau and Hu (2002)
10	Network effect [it only makes sense to use the system if all members of the department use it, otherwise the data is incomplete]	Kralewski et al. (2008)
11	Leaders are not using the system [lack of leadership support/involvement, role model]	Boonstra et al. (2014)
12	Lack of IT infrastructure [insufficient hardware]	Villalba-Mora et al. (2015) Lærum et al. (2001)
13	Insufficient speed / response times [waiting for system responses]	Boonstra et al. (2014) L. Wu, Li, and Fu (2011)
14	Lack of IT support / technical assistance	Heisy-Grove and Patel (2014) Boonstra and Broekhuis (2010) Boonstra et al. (2014)
15	Lack of integration with existing systems [non HIS-systems such as administrative backend etc.]	Villalba-Mora et al. (2015)
16	Lack of user's IT skills	Villalba-Mora et al. (2015) Boonstra and Broekhuis (2010)
17	Lack of knowledge and training on the system	Villalba-Mora et al. (2015) Boonstra et al. (2014)
18	Lack of system reliability [system outages]	Boonstra et al. (2014) L. Wu et al. (2011)

continues on following page

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Continued

	Inhibitor	Source(s)
19	System vendor is not responsive to change requests [resulting in user frustration]	Boonstra et al. (2014)
20	Patient data in the HIS may not be complete [mistrust in data completeness]	Lærum et al. (2001)
21	HIS provides individual's data for performance review [fear of misuse of transparency generated through HIS]	Liu and Cheng (2015)

We also excluded individual physicians working in private and small practice because their behavior is driven by different motives than hospital physicians. For example, if a physician owns her or his own practice, then his or her behavior toward IS is more likely to be driven by cost considerations (Boonstra & Broekhuis, 2010) such as the cost of licensing, implementing and maintaining the information systems. Financial incentives (e.g. those awarded in conjunction with the HITECH act in the US or the introduction of DRG in Germany) also play a significant role in their behavior towards information system (Heisy-Grove & Patel, 2014). However, those physicians working in a hospital are not usually motivated by these considerations.

We collected our data from June to July 2015. In order to capture a broad range of opinions in alignment with the exploratory character of this study, users from different departments, in different roles, and using different systems were selected. In all, we interviewed 48 informants. Their demographics are summarized in Table 2 below.

Data was collected at two hospitals in Germany: Hospital A is a large university medical center, ranking among the 10 largest hospitals in Germany. The hospital comprises of several different clinics which are departments led by a head physician who has both medical and management responsibilities. Hospital A has a very complex and heterogeneous IT landscape. In many cases, the clinics have their own IT departments. Hospital A has a unified administrative back office system (which also provides medical record functionality) complemented with several different HIS systems in different clinics. Several clinics use paper-based medical records or a mix between electronic and paper-based records. Also, there is a variety of different electronic medical records in use at different stages of implementation. Hospital B is a medium-sized specialized clinic which provides a uniform HIS system for all physicians. However, the hospital still relies to a large extend on paper-based medical records.

Table 2. Interview partner demographics

Department	Interviewees	Age		Sex	
Intensive care	11	20-29	9%	Female	48%
Surgery	5	30-39	43%	Male	52%
Internal med	8	40-49	28%		
Orthopedics	13	50-59	17%		
Neurology	8	60+	3%		
Anesthesia	3				
Sum	**48**				

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

3.2. Semi-Structured Interviews

Our semi-structured interview guideline consisted of three sections:

1. **"Who are you and what do you do?":** Demographic data of the interviewee and specification of her/his workplace and/or specific tasks/role(s).
2. **"Do you like the HIS?"** Enablers and inhibitors identified in previous research with special consideration of the inhibiting factors.
3. **"How do you use the HIS in your work?":** Description of the way the interviewee uses the system as part of her/his daily working routines.

The interviewers collected complete demographic information (Section 1) either directly from the informant or through secondary sources. Section 2 was an open discussion and the interviewer drew specific attention to the inhibitors listed above. Given the open question format to address issues, interviewees often picked the most pressing issues and talked about these for some time. Towards the last third of the interview, the interviewer addressed Section 3 to gain an understanding of the way the physician uses the HIS to perform her/his daily tasks.

3.3. Workplace Shadowing

A common problem in social sciences is the disparity between self-reported behavior and actual observable actions. To mitigate this effect, we chose to not only rely on interviews, but rather to also shadow physicians. The team spent over 40 hours accompanying physicians on their ward rounds, observing their general work and team meetings. Our main goal was to observe directly how the physician interacts with the system.

3.4. Data Analysis

We interpreted the data collected as recommended by Miles, Huberman, and Saldana (2013): All notes taken during the interviews were reviewed, clarified as necessary and coded. Open coding was guided by associating the key phrases to either one of the known enablers/inhibitors (section 2.4), or by associating it to a new concept. Following open coding, axial coding was performed to ensure all important aspects had been identified. Two researchers coded the data independently and all disputes were discussed until a unanimous agreement was reached.

The resulting data table lists the major categories and associated concepts, which was the basis for developing a framework to structure the findings.

4. DISCUSSION OF FINDINGS

Our findings underscore that physicians are not generally technophobe, as we had suspected at the outset. They value technology but not for the sake of technology itself but on a much more rational level (in line with Lowenhaupt (2004): Technology is valued most when it is useful to complete a task or process in a more efficient or effective way compared to a given alternative. This "alternative" is usually not a

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

different system, but a workaround (e.g. writing paper notes instead of entering the data directly into the system) or delegating interaction with the system (e.g. by asking a team member to enter data) (cf. Ash et al. (2004).

We found evidence which supports both (semi-)rational and emotional arguments for the way physicians interact with the provided HIS. The main categories resulting from our coding are illustrated in Figure 1 below.

Our interpretation of the interview data and our observations indicate a clear relationship between user, process and system - all influencing each other respectively. Leadership and organization also play an important role. These categories and their relationships are discussed in the following sections.

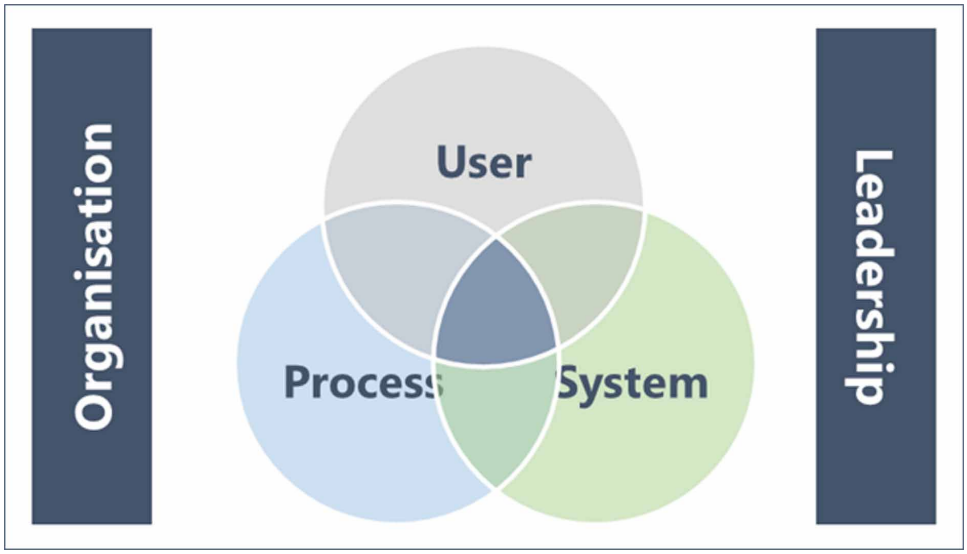
4.1. User - Process – System

Our sample represents a wide range of HIS users: those who are happy with the HIS and those who openly dislike it. In line with the findings of Chau and Hu (2002), we see that physicians usually show a positive attitude towards the HIS when it closely matches their established work behavior, i.e. they perceive a good fit between user, system and process. Also, we noticed, that physicians' attitudes are strongly influenced by the mindset and behavior of their superiors (see Section 4.2).

4.1.1. User

The issues in the category "user" relate predominantly to knowledge and skills, but also to personal efficiency and professional autonomy. We observed that the users were happy with the HIS when they believed the HIS would enable them to get their work done faster or better (i.e. be more efficient and/or effective). The principle enablers were better exchange of/access to information, being able to read all entries [issue of bad handwriting on paper files] and decision support. Although the informants had

Figure 1. Framework of Findings



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

different attitudes towards IT in general, no one refused to work with the system. Some felt that their work is increasingly becoming too IT-focused ("At some point in time we all will have to study computer sciences to do our job…" [N04]).

The physicians interviewed assessed the impact of the HIS on the patient-physician relationship differently depending in part on her/his specialization. Physicians who have more physical interaction with the patient (e.g. in orthopedics) considered the HIS less helpful. In general, physicians appreciated patients' positive perceptions of their use of advanced technology ("I like working with the system. It shows our patients that we are up to date… [with technology]" [N05]). Several informants complained about the HIS lagging behind hardware and software developments like smartphones, tablets and apps. A general perception was that "My work as a doctor is not valued when the IT I am given does not meet my needs…" [P19]. This ties into other recent research, which indicates that physicians tend to perceive IT gadgets as status symbols (Weber, Gewald, Weeger, & Haase, 2016).

The effect of training was considerable. Physicians who attended system training sessions were more comfortable using the system and observably faster and more knowledgeable about the system's functionality. Physicians acknowledge that training is helpful ("It was very hard to use the system in the beginning, but the training helped a lot to make the most out of it…" [P07]), however previous research also observes a strong tendency for physicians to not attend training sessions (Weber & Gewald, 2014).

Several physicians also identified over-transparency (e.g. with regard to performance reviews or lawsuits) stemming from built-in activity tracking and timestamps as a potential inhibitor.

4.1.2. Process

Our findings in the category "process" focus on the intersection of healthcare delivery and HIS use (diagnosis, consultation, prescriptions, etc.)

Our findings underscore that HIS has the greatest potential to benefit physicians if used as designed. When information is entered in a timely and accurate fashion, some important tasks, such as writing the doctor's letter, can be completed very quickly: "We continuously update our doctor's letters as part of our documentation and so in the end they are done very quickly." [P13]. However, this also requires all necessary systems to be integrated: "Writing doctor's letters is the task we like to postpone most because it takes so much time to retrieve all information needed from the system." [P04]. However, some informants mentioned information overload ("I don't want all the information automatically thrown at me. As a doctor, I want to think and make my own decisions." [P10] and others have a general mistrust in the information provided by the system ("With all that copying and pasting one cannot trust the information in the system all the time." [P14]). Hence, system integration and data integrity seem to be a key requirement for successful working processes.

Established work habits can be a major inhibitor to system use. We observed that physicians are rarely willing to adapt their working procedures to fit the system. They rather want the system to reflect their individual (or departmental) working habits. This poses a problem to off-the-shelf software which typically has a limited range of customizability. Even when software can be customized, history (especially in the manufacturing sector, see SAP implementations in the mid 1980ies) has shown that heavy customization leads to system incompatibilities and update problems over time. Individually developed software would resolve this problem, but high maintenance costs are prohibitive for most hospitals.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

System: A major "system" critique associated with the actual HIS and its deployment was that the graphical user interface (GUI) was frequently regarded as too complex or "*not made for doctors*" [P05]. This ties into the "process" findings about established working habits. Some physicians also complained about multiple logons to different systems (as opposed to a single sign on) or the need to press too many buttons before being able to retrieve information (Observation [P02]).

A recurring infrastructure critique was long response times (subjective observation during workplace shadowing) and system outages/crashes. Although concerns with IT security and data privacy were frequently reported in previous research, these issues were not mentioned by our informants.

4.1.3. Interrelationships Among User, Process and System

These three categories appear to be tightly coupled. As we know from the concept of Task-Technology-Fit (Goodhue & Thompson, 1995), a system is accepted most readily when the technology provided fits the task the user needs to perform. In the case of our research, we observe that physicians do not oppose the HIS provided per se, but struggle with the GUI, (perceived) long response times and incomplete or not easily accessible information. Some of these factors can be addressed by training and investment in IT-infrastructure, but an off-the-shelf software almost certainly will not support all established working processes in all hospitals. So in order to support physicians efficiently, the systems need to change (improve GUI, implement single sign on, improve data integration) the processes need to change (adapt clinical processes to a standard model which can be supported by a commercial software) and the users need to change (attend training sessions).

These three factors are so closely interconnected that addressing just one of them appears to be an inviable path forward.

4.2. Leadership

Our interviews and observations also offer important "leadership" insights into how medical and administrative managers influence the user-process-system framework. In a hospital context, we distinguish between medical management (leadership exerted by the medical superiors, e.g. the head physician) and administrative management, i.e. the hospital administration.

4.2.1. Medical Management

In hospitals, the medical management generally fulfills the administrative and the medical management role because the head physician bears overall responsibility for her/his department, including medical, organizational and financial responsibility. This distinguishes the healthcare context from other industries where managers often do not have direct operational responsibility. In hospitals, the head of the surgery department usually carries out most difficult procedures personally, whereas the head of engineering typically does not maintain engines in the automotive industry. This has important implications for research. Our findings clearly substantiate the important role superiors play in forming attitudes and behavior towards HIS in their departments. Whenever the head physician was skeptical of HIS, e.g. "Technology is necessary these days but paper is faster and more efficient [P01]", this attitude was common across the department. This applies to positive mindsets as well. The strong impact of medical management

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

towards HIS derives from the double role of medical and managerial responsibility. The department head greatly influences working procedures, has a (if not THE) leading role in the HIS selection process, and decides how much budget will be spent on system development, such as new functionalities and integration with other systems.

4.2.2. Administrative Management

Hospital administration traditionally has a difficult role to play vis-à-vis medical directors because they are financially responsible for the hospital but have no say medically. This makes it difficult to assert overall leadership in the absence of direct supervision. As a result of the lack of supervision, different departments sometimes deploy different HIS for the same purpose. This is costly and greatly hinders integration and standardization efforts. However, it is difficult for the hospital administration to influence individual departments' selection of an HIS if the selection is made on the basis of medical procedures.

4.2.3. Influence of Leadership on User-Process-System

Our findings show that leadership on all levels plays an important role. For example, in units where the medical management (head physician) is a strong supporter of HIS, the attitude of the physicians in the unit towards the system was much more positive than in other units. This is not surprising because the head physician bears ultimate responsibility for all actions across the entire department and thus strongly influences working practices.

When medical management encourages the physicians in his or her unit to use the system, they are much more satisfied with the system. If administrative management (especially together with medical management) fosters consolidation, integration and in some instances standardization of systems, the benefit for all users of the HIS increases. The physicians we interviewed in the hospital with a unified HIS were generally more satisfied with the system than the physicians we interviewed where the IT landscape was diverse.

If management has a shared vision of HIS use and makes an effort to make this vision a practical reality, the overall outcome is rewarding because it directly influences the user-process-system construct.

4.3. Organization

The main issue raised by the physicians related to the "organization" of work within the hospital was the low level of integration with other departments. Due to the division of administrative and medical leadership described above, the hospital lacked a coherent system and process integration structure, resulting in a complex IT landscape consisting of several different HIS systems and heterogeneous working procedures: "I have to put the request for a counsel into the system, print and sign it, fax the request to the other department and then call them every time to make sure they got the information" [P22].

Many physicians also commented that the IT department was too distant from the medical profession and that they found it difficult to communicate and interact with IT personnel. This observation relates to previous IS research into IT/business alignment which shows that hospitals with good alignment perform better than hospitals with isolated departments (Weeger, 2017).

As illustrated, the organizational factor indirectly but strongly impacts the user-process-system construct. Aligning the organization well leads to overall performance improvements.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4.4. Summary

In summary, our interviews and workplace shadowing showed that physicians are willing to work with systems when these systems help them do their job more efficiently. We observe that the better the user-process-system construct is aligned, the better the physicians interact with the system and the better the results are. We also observed that strong and coordinated leadership enables good use of IT and that organizational alignment positively influences system use.

Our resulting framework consists of an inner model (user-process-system) of cross-influential factors and an outer model (leadership and organization) which directly influence the inner model.

5. IMPLICATIONS

Our findings show that different IS research strands are required to explain use of a mandatory HIS by hospital physicians, including Task-Technology-Fit and business/IT alignment. We do not consider our understanding of the issues to be refined enough to merit large-scale quantitative research assessing enablers and inhibitors and recommend further research building on our framework which certainly needs to be refined when deeper understanding is available.

From a practical perspective, our study identified several issues frequently overlooked in HIS system implementation. As in other industries, hospitals are in need of Business Process Reengineering. Physicians' use of a HIS is less likely to be improved by buying a (new) system than through a combined effort of system customization and adapted working habits, which requires significant change management effort. Additionally, more focus should be put on GUI development and IT topics which directly impact users, such as response times and multiple logon requirements. Although it may never be possible to satisfy all users, such efforts can increase user satisfaction.

6. LIMITATIONS AND FURTHER RESEARCH

This qualitative study provides deeper insights into hospital physicians' use of mandatory HIS system and aims to identify major use inhibitors. Since we only interviewed 48 physicians at 2 hospitals over a short timeframe in Germany, our results are not widely generalizable.

However, future research can nonetheless benefit from our research by identifying a basis for comparisons across different hospitals, organizational structures, etc. We encourage IS usage scholars to test our findings in other settings.

7. CONCLUSION

We conducted an exploratory study into the major inhibitors of hospital medical personnel using the mandatory HIS system(s) the hospital provides. Our findings show that in addition to the traditional focus areas of information systems research "user, process and system", leadership and organizational setting also influence user behavior. We find that despite a generally positive attitude towards the system on the part of physicians, effective use was inhibited primarily by insufficient support of the working procedures caused by:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. Insufficient systems integration and/or inadequate systems deployed by the hospital,
2. Traditional work processes that do not match the systems in place, and
3. Physicians' reluctance to adapt their work processes.

ACKNOWLEDGMENT

The authors presented an earlier version of this paper on the 50th Hawaii International Conference on System Sciences (HICSS), Big Island, HI, 2017.

REFERENCES

Abouzahra, M., Guenter, D., & Tan, J. (2015). Integrating Information Systems and Healthcare Research to Understand Physicians' use of Health Information Systems: a Literature Review. *Paper presented at the International Conference on Information Systems*, Dallas, TX.

Ash, J. S., Berg, M., & Coiera, E. (2004). Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. *Journal of the American Medical Informatics Association*, *11*(2), 104–112. doi:10.1197/jamia.M1471 PMID:14633936

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *Management Information Systems Quarterly*, *25*(3), 351–370. doi:10.2307/3250921

Bhattacherjee, A., & Hikmet, N. (2007). Physicians' resistance toward healthcare information technology: A theoretical model and empirical test. *European Journal of Information Systems*, *16*(6), 725–737. doi:10.1057/palgrave.ejis.3000717

Boonstra, A., & Broekhuis, M. (2010). Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. *BMC Health Services Research*, *10*(1), 231. doi:10.1186/1472-6963-10-231 PMID:20691097

Boonstra, A., Versluis, A., & Vos, J. F. (2014). Implementing electronic health records in hospitals: A systematic literature review. *BMC Health Services Research*, *14*(370), 1–24. PMID:25190184

Brown, S. A., Massey, A. P., Montoya-Weiss, M. M., & Burkman, J. R. (2002). Do I really have to? User acceptance of mandated technology. *European Journal of Information Systems*, *11*(4), 283–295. doi:10.1057/palgrave.ejis.3000438

Buddrus, U. (2013). *IT Benchmarking im Krankenhaus mit HIMSS' europäischen EMR Adoption Model*. Retrieved from http://www.ehealthsummit.at/sites/default/files/ehsat_EMRAM%20Workshop_Buddrus.pdf

Buntin, M. B., Burke, M. F., Hoaglin, M. C., & Blumenthal, D. (2011). The benefits of health information technology: A review of the recent literature shows predominantly positive results. *Health Affairs*, *30*(3), 464–471. doi:10.1377/hlthaff.2011.0178 PMID:21383365

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Burton-Jones, A., & Grange, C. (2012). From use to effective use: A representation theory perspective. *Information Systems Research*, *24*(3), 632–658. doi:10.1287/isre.1120.0444

Chan, F. K., Thong, J. Y., Venkatesh, V., Brown, S. A., Hu, P. J., & Tam, K. Y. (2010). Modeling citizen satisfaction with mandatory adoption of an e-government technology. *Journal of the Association for Information Systems*, *11*(10), 519–549.

Chau, P. Y. K., & Hu, P. J.-H. (2002). Investigating healthcare professionals' decisions to accept telemedicine technology: An empirical test of competing theories. *Information & Management*, *39*(4), 297–311. doi:10.1016/S0378-7206(01)00098-2

Davino, J. M. (2011). Complexities of Delivering Health Care. *The Journal of Global Health Care Systems*, *1*(4), 1–9.

DeLone, W. H., & McLean, E. R. (2002). Information systems success revisited. *Paper presented at the 35th Hawaiian International Conference on System Sciences*, Hawaii. 10.1109/HICSS.2002.994345

Denis, P. (2009). Integrated Care Needs Integrated Information Management and Technology. *Healthcare Quarterly, 13*(Sp), 24-29.

Doolin, B. (2004). Power and resistance in the implementation of a medical management information system. *Information Systems Journal*, *14*(4), 343–362. doi:10.1111/j.1365-2575.2004.00176.x

Engelbert, R., & Graeml, A. (2015). Beyond IT Acceptance. *Paper presented at the 21st Americas Conference on Information Systems*, Puerto Rico.

European Commission. (2013). *Anteil der IT-Aufwendungen am Gesamtbudget europäischer Krankenhäuser im Jahr 2013*. Retrieved from http://de.statista.com/statistik/daten/studie/423129/umfrage/anteil-von-it-aufwendugen-am-gesamtbudget-europaeischer-krankenhaeuser/

Gallivan, M. J. (2001). Organizational adoption and assimilation of complex technological innovations: Development and application of a new framework. *The Data Base for Advances in Information Systems*, *32*(3), 51–85. doi:10.1145/506724.506729

Gebauer, L., Söllner, M., & Leimeister, J. M. (2013). Towards understanding the formation of continuous IT use. *Paper presented at the 34th International Conference on Information Systems*, Milano.

Gold, M. R., McLaughlin, C. G., Devers, K. J., Berenson, R. A., & Bovbjerg, R. R. (2012). Obtaining providers' 'buy-in'and establishing effective means of information exchange will be critical to HITECH's success. *Health Affairs*, *31*(3), 514–526. doi:10.1377/hlthaff.2011.0753 PMID:22392662

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *Management Information Systems Quarterly*, *19*(2), 213–236. doi:10.2307/249689

Heisey-Grove, D., & Patel, V. (2014). *Physician motivations for adoption of electronic health records*. Washington, DC: Office of the National Coordinator for Health Information Technology.

Hirschheim, R., & Newman, M. (1988). Information systems and user resistance: Theory and practice. *The Computer Journal*, *31*(5), 398–408.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Holden, R. J., & Karsh, B.-T. (2010). The Technology Acceptance Model: Its past and its future in health care. *Journal of Biomedical Informatics*, *43*(1), 159–172. doi:10.1016/j.jbi.2009.07.002 PMID:19615467

IDC. (2015). *IDC Health Insights Western European Healthcare Providers IT Spending Forecast 2014–2018* (#HIOH53W).

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *Management Information Systems Quarterly*, 23(2), 183–213.

Kellermann, A. L., & Jones, S. S. (2013). What it will take to achieve the as-yet-unfulfilled promises of health information technology? *Health Affairs*, *32*(1), 63–68. doi:10.1377/hlthaff.2012.0693 PMID:23297272

Kim, H.-W., & Kankanhalli, A. (2009). Investigating user resistance to information systems implementation: A status quo bias perspective. *Management Information Systems Quarterly*, *33*(3), 567–582.

Koppel, R. (2016). Great Promises of Healthcare Information Technology Deliver Less. In C. A. Weaver, M. J. Ball, G. R. Kim, & J. M. Kiel (Eds.), *Healthcare Information Management Systems* (pp. 101–125). Springer International Publishing. doi:10.1007/978-3-319-20765-0_6

Kralewski, J. E., Dowd, B. E., Cole-Adeniyi, T., Gans, D., Malakar, L., & Elson, B. (2008). Factors influencing physician use of clinical electronic information technologies after adoption by their medical group practices. *Health Care Management Review*, *33*(4), 361–367. doi:10.1097/01.HCM.0000318773.67395.ce PMID:18815501

Lærum, H., Ellingsen, G., & Faxvaag, A. (2001). Doctors' use of electronic medical records systems in hospitals: Cross sectional survey. *BMJ (Clinical Research Ed.)*, *323*(7325), 1344–1348. doi:10.1136/bmj.323.7325.1344 PMID:11739222

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *Management Information Systems Quarterly*, *29*(3), 461–491.

Laumer, S., & Eckhardt, A. (2012). Why Do People Reject Technologies: A Review of User Resistance Theories. In Y. Dwivedi, M. Wade, & S. Schneberger (Eds.), *Information Systems Theory. Integrated Series in Information Systems* (Vol. 28). New York, NY: Springer.

Laumer, S., Maier, C., Eckhardt, A., & Weitzel, T. (2014). Why are they grumbling about my new system? Theoretical foundation and empirical evidence of employee grumbling as a user resistance behavior. *Paper presented at the 35th International Conference on Information Systems*, Auckland.

Laumer, S., Maier, C., Eckhardt, A., & Weitzel, T. (2015). User personality and resistance to mandatory information systems in organizations: A theoretical model and empirical test of dispositional resistance to change. *Journal of Information Technology*.

Limayem, M., Hirt, S. G., & Cheung, C. M. (2007). How habit limits the predictive power of intention: The case of information systems continuance. *Management Information Systems Quarterly*, *31*(4), 705–737.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, C.-F., & Cheng, T.-J. (2015). Exploring critical factors influencing physicians' acceptance of mobile electronic medical records based on the dual-factor model: A validation in Taiwan. *BMC Medical Informatics and Decision Making*, *15*(1), 125. doi:10.118612911-014-0125-3 PMID:25889506

Lowenhaupt, M. (2004). Removing barriers to technology. *Physician Executive*, *30*(2), 12–14. PMID:15085666

Melone, N. P. (1990). A theoretical assessment of the user-satisfaction construct in information systems research. *Management Science*, *36*(1), 76–91. doi:10.1287/mnsc.36.1.76

Miles, M. B., Huberman, A. M., & Saldana, J. (2013). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). Sage.

Paré, G., Sicotte, C., & Jacques, H. (2006). The effects of creating psychological ownership on physicians' acceptance of clinical information systems. *Journal of the American Medical Informatics Association*, *13*(2), 197–205. doi:10.1197/jamia.M1930 PMID:16357351

Shachak, A., & Reis, S. (2009). The impact of electronic medical records on patient–doctor communication during consultation: A narrative literature review. *Journal of Evaluation in Clinical Practice*, *15*(4), 641–649. doi:10.1111/j.1365-2753.2008.01065.x PMID:19522722

Shaikh, A. A., & Karjaluoto, H. (2015). Making the most of information technology and systems usage: A literature review, framework and future research agenda. *Computers in Human Behavior*, *49*, 541–566. doi:10.1016/j.chb.2015.03.059

Thomas, T. (2008). A Computer Security Expert's Perspective on Electronic Medical Records. *Journal of Controversial Medical Claims*, *15*(2), 19–20.

Tulu, B., Burkhard, R., & Horan, T. A. (2006). Information systems and health care XIV: continuing use of medical information systems by medical professionals: empirical evaluation of a work system model. *Communications of the Association for Information Systems*, *18*(1), 31.

Vehring, N., Riemer, K., & Stefan, K. (2011). "Don't pressure me!" Exploring the Anatomy of Voluntariness in the Organizational Adoption of Network Technologies. *Paper presented at the Proceedings of the 32nd International Conference on Information Systems*, Shanghai.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. doi:10.1287/mnsc.46.2.186.11926

Villalba-Mora, E., Casas, I., Lupiañez-Villanueva, F., & Maghiros, I. (2015). Adoption of health information technologies by physicians for clinical practice: The Andalusian case. *International Journal of Medical Informatics*, *84*(7), 477–485. doi:10.1016/j.ijmedinf.2015.03.002 PMID:25823578

Wakefield, D. S., Halbesleben, J. R., Ward, M. M., Qiu, Q., Brokel, J., & Crandall, D. (2007). Development of a measure of clinical information systems expectations and experiences. *Medical Care*, *45*(9), 884–890. doi:10.1097/MLR.0b013e3180653625 PMID:17712259

Walter, Z., & Lopez, M. S. (2008). Physician acceptance of information technologies: Role of perceived threat to professional autonomy. *Decision Support Systems*, *46*(1), 206–215. doi:10.1016/j.dss.2008.06.004

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Weber, M., Gewald, C., Weeger, A., & Haase, U. (2016). The Effect of 'Device' in Task-Technology Fit A study of German Hospitals. *Paper presented at the Forty-Ninth Hawaii International Conference on Systems Sciences*, Hawaii. 10.1109/HICSS.2016.402

Weber, M., & Gewald, H. (2014). "I don't have time for that" - Examining the Impact of Training on Physician's Task Performance. *Paper presented at the 5th Workshop on Health IT and Economics*, Washington.

Weeger, A. (2017). What Qualifies Project-related Artefacts to Bridge Boundaries in IS Implementation Projects – An Activity Theoretical Perspective. *Paper presented at the 50th Hawaiian International Conference on System Sciences*. 10.24251/HICSS.2017.621

Weeger, A., & Gewald, H. (2015). Acceptance and use of electronic medical records: An exploratory study of hospital physicians' salient beliefs about HIT systems. *Health Systems (Basingstoke, England)*, *4*(1), 64–81.

Wiggins, C., Peterson, T., & Moss, C. (2015). Ambulatory surgery centers′ use of Health Information Technology. *Health Policy and Technology*, *4*(2), 100–106. doi:10.1016/j.hlpt.2015.02.006

Wu, J., & Du, H. (2012). Toward a better understanding of behavioral intention and system usage constructs. *European Journal of Information Systems*, *21*(6), 680–698. doi:10.1057/ejis.2012.15

Wu, L., Li, J.-Y., & Fu, C.-Y. (2011). The adoption of mobile healthcare by hospital's professionals: An integrative perspective. *Decision Support Systems*, *51*(3), 587–596. doi:10.1016/j.dss.2011.03.003

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 13(1); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 29-44; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 87
Using Key Performance Indicators to Reduce Perceived Perioperative Complexity and Improve Patient Workflow

Jim Ryan
Auburn University at Montgomery, USA

Barbara Doster
University of Alabama, USA

Sandra Daily
Duke Medical Center, USA

Carmen Lewis
Troy University, USA

ABSTRACT

This study examines the development and use of balanced scorecard metrics as key performance indicators within each stage of the perioperative process to enable business process management across the entire process to gauge performance and target improvement opportunities. The identification of existing limitations, potential capabilities, and the subsequent contextual understanding are contributing factors toward perioperative improvement. This paper identifies how dynamic technological activities of analysis, evaluation, and synthesis applied to internal and external organizational data can highlight complex relationships within integrated hospital processes to address root causes rather than symptoms and ultimately yield improved capabilities. This case study investigates how integrated information systems can identify, qualify, and quantify perioperative performance indicators to measure improvement based on a 157-month longitudinal study of a large, 1,157 registered-bed teaching hospital. The theoretical and practical implications and/or limitations of this study's results are also discussed with respect to practitioners and researchers alike.

DOI: 10.4018/978-1-7998-1204-3.ch087

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The current focus of healthcare administration in the United States has shifted toward performance monitoring and improving clinical outcomes to meet regulatory and reimbursement requirements due to the American Recovery and Reinvestment Act of 2009, the Health Information Technology for Economic and Clinical Health Act, the Affordable Care Act, and the Joint Commission on Accreditation of Healthcare Organizations (TJC) / Centers for Medicare & Medicaid Services (CMS) core measures (Blumenthal, 2012). Meeting these performance and reporting challenges require leveraging information systems (IS) and technologies (IT) (PwC, 2012). Consequently, the current widespread IS/IT adoption across United States healthcare also necessitates the need for value realization (Jones et al., 2014). To this end, this study investigates how key performance indicators (KPIs) across a hospital's entire perioperative process can identify and ultimately achieve improved patient workflow via business process management (BPM).

A hospital's perioperative process is complex (Fowler et al., 2008), involving multiple interconnected sub-processes that provide surgical care for inpatients and outpatients during pre-assessment, pre-operative, intra-operative, post-operative, and central sterile supply (CSS) activities. Furthermore, the perioperative process yields patient end-state goals (Silverman & Rosenbaum, 2009) where: (1) a correct diagnosis for surgical intervention is identified with noted co-morbidities and patient consent; (2) a patient undergoes the surgical procedure; (3) a patient exhibits minimal exacerbation of existing disorders; (4) a patient avoids new morbidities; and (5) a patient experiences prompt procedure recovery. Meeting perioperative patient end-state goals also avoids hospital-acquired-conditions and hospital-acquired-infections (HACs/HAIs) connected with negative financial incentives (CMS, 2015; Waters et al., 2015; NQF, 2008). However, perioperative sub-processes are sequential, where sub-process activity sequence paces the efficiency and effectiveness of subsequent sub-process activities and risks achievement of patient end-state goals. Consequently, the perioperative process is tightly coupled to patient safety, patient quality of care, patient flow, and stakeholders' satisfaction (i.e., patient, physician/surgeon, nurse, perioperative staff, and hospital administration).

This case study identifies complex perioperative dynamics within and across the entire process that is nested in the hospital environment and reflected via sub-process scorecard metrics. This research investigates how traditional BPM practices are applicable to reduce perioperative complexity, measure process performance, target areas for improvement, and subsequently yield improved patient workflow. This study associates particular perioperative sub-processes with individual patient end-state goals, specifically highlights multiple balanced scorecard metrics from each perioperative sub-process, and explains how the metrics are applicable as KPIs. The investigation method covers a longitudinal study of an integrated clinical scheduling information system (CSIS) within an academic medical center. The implementation of agile, integrated IS and subsequent contextual understanding of perioperative sub-process data prescribed opportunities for measured improvements. Specifically, developing and implementing perioperative sub-process metrics as KPIs, grounded in internal and external best-practices, provides the framework for targeting opportunities and evoking improvement. The combined assessment of sub-process KPIs also provide change dynamics for evaluation and improvement to the overall perioperative process.

The following sections review previous literature on KPIs, BPM, as well as perioperative scorecards and dashboards. Following the literature review, we present our methodology, case study background, observed effects, and summary discussion. By identifying a holistic framework for analysis, evaluation, and synthesis of sub-process KPIs with established benchmarks, this paper prescribes an a priori envi-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ronment to support perioperative sub-process measurement and improvement through the identification of specific sub-process root causes rather than overall process symptoms. The conclusion also addresses study implications and limitations.

LITERATURE REVIEW

Integrated IS offer continuity through information sharing, synergy, and improvement (Karimi, 1988). Likewise, integrated IS and IT provide measurement and subsequent accountability for healthcare quality and cost, creating a dichotomy (e.g., quality versus cost) that represents the foundation for healthcare improvement (Dougherty & Conway, 2008). Within the perioperative process, patient end-state goals are the focus of work. However, United States hospitals currently face increasing pressure to provide objective evidence of organizational quality, efficiency, and effectiveness (Blumenthal, 2012; PwC, 2012). To this end, a BPM approach (Jeston & Nelis, 2008; Turban et al., 2008) borrowed from the manufacturing industry provides a framework to measure performance and target improvement.

The perioperative process is typically the primary source of hospital admissions, averaging between 55 to 65 percent of overall hospital margins (Peters & Blasco, 2004). Likewise, total hospital supply costs are variable with the largest cost category being the perioperative process (e.g., 33 percent) (Macario et al., 1995). Nonetheless, the perioperative process has many core sub-processes nested within the hospital environment that yield overall clinical performance. Furthermore, managing and optimizing quality, efficient, and cost-effective perioperative sub-processes to yield improved clinical outcomes are critical success factors (CSFs), both operationally and financially. To meet these demands, administrators and medical professionals must focus technology-enhanced practices that yield high quality of care and patient safety, coupled with increased efficiency and cost effectiveness. Measured utilization of these practices is not a result from any lack of research as an extensive body of knowledge exists concerning the application of these approaches in healthcare (Doebbeling et al., 2012; Albanese et al., 2010; Eckerson, 2010; and Tarantino, 2003). However, the literature suggests that such management practices and interventions can yield positive results with significant variations in success.

Key Performance Indicators

Early in the IT literature, Ackoff (1967) proposed IS design to include feedback control to avoid management misinformation. Similarly, information before and after intervention is an integral part of process improvement, which encourages bottom-up communication at the day-to-day operations level and requires process data comparisons to control metrics. Consequently, organizations define data metrics as KPIs to assist management in monitoring CSFs for organizational action via business processes (Munroe & Wheeler, 1980; Rockart, 1979; Zani, 1970). To this end, KPIs focus on organizational priorities that create long-term value for stakeholders (Baker, 2015). Likewise, performance metrics are essential requirements for BPM. However, doubt exists as to whether perioperative management can meet these demands (Catalona & Fickenscher, 2007), in part due to perceived perioperative complexity as a barrier (Fowler et al., 2008).

Operational and tactical KPIs identified in evaluating a hospital's perioperative process typically focus on intra-operative metrics (Doebbeling et al., 2012; Wright et al., 2010; Peters & Blasco 2004; Tarantino, 2003; and Kanich & Byrd, 1996) from the operating room (OR) like: (1) monitoring the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

percentage of surgical cases that start on-time (OTS), (2) OR turnaround time (TAT) between cases, (3) OR suite utilization (UTIL), and (4) labor hours per patient care hours or units-of-service (UOS) expended. However, OR schedules are tightly coupled to individual OR suites, patients, and surgeons where incomplete perioperative sub-process tasks upstream impact intra-operative KPIs. For example, incomplete pre-assessment or pre-operative tasks as well as unavailable CSS supplies/instruments/devices can delay a scheduled case as well as the subsequent scheduled cases in the particular OR suite or for the particular surgeon. Consequently, intra-operative inefficiencies and delays in turn impact post-operative activities downstream. Within the perioperative sub-processes, symptoms downstream can be more noticeable than actual root causes upstream, which supports having KPIs within all the perioperative sub-processes to gauge performance independently to identify problems and root causes rather than symptoms. Addressing problems and root causes supports continuous improvement.

Perioperative sub-process KPIs are often closely associated with multiple hospital CSFs. For example, OR TAT and a flexible work environment are CSFs for physician satisfaction (Wright et al., 2010; Macario, 2006; and Tarantino, 2003), which in turn is a CSF for hospital margin. Conversely, poor operational and tactical KPIs like OTS and OR TAT affect strategic CSFs of patient safety, patient quality of care, surgeon/staff/patient satisfaction, and hospital margin (Marjamaa et al., 2008; Peters & Blasco, 2004). Likewise, inefficient and ineffective reprocessing of instruments/devices within the CSS sub-process yields poor OR TAT KPIs (Duro, 2015) as well as places strategic CSFs at risk. This study takes a holistic view and identifies KPIs across all the perioperative sub-processes.

Business Process Management

This study uses the BPM definition provided by Jeston and Nelis (2008, p. 10) as "the achievement of an organization's objectives through the improvement, management, and control of essential business processes." The authors further elaborate that process management and analysis is integral to BPM, where there is no finish line for improvement. Hence, this study views BPM as an organizational commitment to consistent and iterative process performance improvement that meets organizational objectives.

Continuous process improvement (CPI) is a systematic approach toward understanding process capability, customer's needs, and sources of observed variation. Tenner and DeToro (1997) views CPI as an organizational response to an acute crisis, a chronic problem, and/or an internal driver. Incremental improvement gains occur via iterative cycles of analysis, evaluation, and synthesis or plan-do-study-act (Walton, 1986) to minimize observed variation. Doubt can exist as to whether the incremental improvement addresses symptoms versus causes; whether the improvement effort is sustainable year after year; and/or whether management is in control of the process (Jeston & Nelis, 2008).

Business analytics is the body of knowledge identified with technology solutions that incorporate performance management, definition and delivery of business metrics, as well as data visualization and data mining (Turban et al., 2008). Business analytics within BPM focus on the effective use of organizational data and information to drive positive business action (Tenner & DeToro, 1997). The effective use of business analytics demands knowledge and skills from subject matter experts and knowledge workers. Similarly, Wears and Berg (2005) concur that IS and/or IT only yield high-quality healthcare when the use patterns are tailored to knowledge workers and their environment.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Perioperative Scorecards and Dashboards

The terms scorecard and dashboard are frequently used synonymously to describe a visual display mechanism that conveys performance information at a glance (Baker, 2015). Dashboards and scorecards are comprised almost entirely of KPIs, which collectively embody the corporate strategy tailored to each individual in the organization by role and level (Eckerson, 2006). Using scorecards and dashboards for purposes such as strategic, tactical, and day-to-day operations, coupled with internal and external best-practice benchmarks, provide the framework for targeting opportunities and measuring improvement. Drawing a distinction between the two, dashboards monitor process performance via charts or tables while scorecards focus on target or goal attainment using graphics (Baker, 2015; Wilbanks & Langford, 2014).

The KPI data presented in scorecards or dashboards provides the opportunity to monitor performance for process improvement or gather data for timely information in support of decision making. KPI data granularity within the scorecard or dashboard must be high enough to support drill down capabilities for knowledge discovery via online analytical processing or data mining to support analysis, evaluation, and synthesis activities associated with possible decisions (Karami, 2013). Overall, scorecards and dashboards can play a pivotal role in converting data and information into actionable knowledge (Baker, 2015; Wilbanks & Langford, 2014; Eckerson, 2006).

RESEARCH METHODOLOGY

The objective of this study is to examine KPIs across all perioperative sub-processes for use in traditional BPM practices that provides a framework to target and measure improvement as well as maintain perioperative patient end-state goals. To this end, case research is particularly appropriate (Weber, 2004; Eisenhardt, 1989). Paré (2001) recommended using a positivist case study methodology to build and test theories in IS research. A positivist approach in case research allows focus and analysis of the associated qualitative problems and environmental complexity (Yin, 2003). Hence, our study took an in-depth positivist approach to case research.

Our research site (e.g. University Hospital) is an academic medical center, licensed for 1,157 beds and located in the southeastern United States. University Hospital is a Level 1 Trauma Center, with a robotics program across eight surgical specialties as well as a Women's/Infant facility. University Hospital's recognition includes Magnet since 2002 and a Top 100 Hospital by U.S. News and World Report since 2005. Concentrating on one research site facilitated the research investigation and allowed collection of longitudinal data. This research spans activities from August 2003 through September 2016, with particular historical data since 1993. During the 157-month study, we conducted field research and collected data via multiple sources including interviews, field surveys, site observations, field notes, archival records, and document reviews.

CASE BACKGROUND

University Hospital Perioperative Services (UHPS) is the University Hospital department designated to coordinate and manage perioperative patient care across pre-admissions via the pre-operative assessment consultation and treatment (PACT) clinic, admissions, surgical preparations (PRE-OP), central sterile

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

supply (CSS), OR surgery and endoscopy, and post anesthesia care units (PACU). The workflow through CSS reprocesses all reusable surgical instruments/devices and moves supplies to pre-assessment, pre-operative, intra-operative, and post-operative activities. The following sections highlight tools, events, and outcomes that have shaped the UHPS BPM approach.

CSIS Implementation

UHPS implemented a new, agile CSIS in 2003, after using its prior CSIS for 10 years. The new CSIS supports OLAP tools, a proprietary structured query language, and both operational and managerial data stores (e.g., operational data store and separate data mart). Flexible routing templates or surgical preference cards (SPCs) allow standardization of surgical care supplies-instruments-devices or customization for surgeon specific procedures. Since the CSIS implementation, over 7,750 generic and custom SPC configurations facilitate the surgical specialty services (SSS) represented in Table 1. Similarly, the agile CSIS data mart serves as the central repository for perioperative data used to report KPIs with a business intelligence layer to support data visualization and dashboards.

November 2004

University Hospital opened a new diagnostic and surgical facility (i.e., North Pavilion) in November 2004. UHPS relocated CSS onto one floor (i.e., 3rd) with Pre-OP, ORs, and PACU on each of the two floors above CSS (i.e., 5th and 7th). The North Pavilion campus expanded UHPS to cover an additional floor and nine ORs (i.e., 33% capacity increase), providing 40 state-of-the-art OR suites with standard-

Table 1. Current CSIS SPCs

Surgical Specialty Service	SPCs
BURN – Trauma burns	26
CARDIO –Cardiovascular & Thoracic	946
ENT – Ear, Nose, & Throat	1,030
GI – Gastro-intestinal	460
GYN – Obstetrics, oncology, incontinence	611
NEURO – Neurological	763
ORAL - Oral Maxilla Facial	236
ORTHO – Orthopedic, joint/device	1,208
PLAS – Plastic surgery	681
SURG ONC – Surgical oncology	329
TX – Transplants (liver, renal)	194
TRAUMA – Trauma, MASH	203
URO – Urology	533
VASCULAR – arteries & blood vessels	558

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ized and surgical specialty equipment. Within six weeks of occupancy, an intra-operative KPI reflected chaos. Surgical case OTS plunged to 18% during December 2004. Having only 18% OTS is unacceptable in a highly competitive hospital industry, as 82% of scheduled surgeries experience delays and risk patient care and safety.

In January 2005, UHPS expressed concerns before a quickly convened meeting of C-level officers with top nursing, surgeons, and anesthesia representatives. The meeting yielded a hybrid management structure and governance in the formation of a multidisciplinary executive team, chartered and empowered to evoke change. The executive team consisted of perioperative stakeholders (i.e., surgeons, anesthesiologists, nurses, and UHPS). The executive team's charter was to focus on patient care and safety, attack difficult questions, and remove inefficiencies. No issue was off-limits.

Perioperative Process Improvements and Expansion

University Hospital's executive team launched a process improvement effort in 2005 to address the perioperative crisis. This CPI effort resulted in the executive team enlisting numerous task forces to address specific perioperative process problems and opportunities, which became the foundation for their BPM approach (Ryan et al., 2008). Since 2005, UHPS has focused data-driven analysis of KPIs to gauge process variance, identify improvement opportunities, and yield process improvement (Ryan et al., 2014). Using this systematic BPM approach, University Hospital's improvement efforts have targeted various perioperative sub-processes to achieve a balanced workflow (Ryan et al., 2016) at the North Pavilion campus. The results achieved through UHPS's BPM approach encouraged University Hospital administration to expand UHPS management responsibilities into other University Hospital facilities.

In 2009, UHPS expanded its management beyond the initial 32 general ORs (GENOR) and 8 cardiovascular OR suites (CVOR) within the North Pavilion campus to the other campuses of University Hospital Health System (UHHS) facilities to include 16 OR suites at the Highland campus (HHOR) and 8 endoscopy labs at the TK Clinic campus. In 2011, UHPS also created and assumed management of a pre-operative assessment, consultation, and treatment (PACT) clinic to screen pre-operative patient flow into UHHS ORs (Ryan et al., 2012). Two additional general ORs have been equipped since 2013 at the North Pavilion campus to bring UHPS management to 58 ORs and 11 endoscopy labs.

OBSERVED EFFECTS

The identification and definition of perioperative KPIs has been an iterative evolution since 2005. Consistently, the focus was data-driven improvements, with process control measures collected through the CSIS and benchmarked to external industry standards or prior months' metrics. While reviewing what could have been done better early in the CPI efforts, UHPS recognized the need to involve more perioperative stakeholders in improvement efforts and not just end-result to-do lists. Consequently, the executive team launched an initiative in 2008 to begin performance reporting for CMS and TJC. The perioperative BPM effort established balanced scorecard (Kaplan & Norton, 1996) measures (e.g. quality of process, satisfaction of customer, or financial) and a means to disseminate the process feedback to perioperative stakeholders at strategic, tactical, and operational levels. The BPM approach was expanded in FY2010 to reinforce UHHS strategy across all core hospital processes (Ryan et al., 2014).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In FY2014, the BPM task force (e.g., surgeons, anesthesiologists, nurses, and UHPS administration) was charged to create actionable information to support operational, tactical, or strategic initiatives across the perioperative process. The cross-disciplinary effort identified KPIs upstream and downstream of the ORs to include all perioperative sub-processes. The following sections detail the resulting KPIs, with each KPI classified by a Kaplan and Norton (1996) balanced scorecard (e.g. BSC) metric category of F-financial, Q-quality of process, or S-satisfaction of customer. A sampling of sub-process KPIs also illustrate the CSIS business intelligence layer capabilities for data visualization and performance dashboards.

Pre-Assessment KPIs

Pre-operative integrated evaluations communicate and document practitioner-patient awareness to avoid conflicts and identify potential OR specific risks (Silverman & Rosenbaum, 2009). To this end, University Hospital's PACT clinic manages pre-admission patient flow into UHHS ORs where all surgical patients receive a scheduled PACT evaluation prior to their surgical procedure. Prior to implementing the PACT clinic in 2011, missed pre-assessment activities contributed to 18% of intra-operative OTS delays.

The pre-assessment sub-process assists with a patient's end-state goal of (1) identifying a correct diagnosis for surgical intervention with noted co-morbidities and patient consent. When a patient arrives for a PACT evaluation, the surgeon documents a focused assessment of the patient and confirms surgery consent. The surgeon may also order testing (i.e. Stress, EKG, Imaging/X-ray, and/or Lab) as well as a cardiac/medical consultation as needed. Documentation of existing HACs or HAIs prior to the patient's admission also disassociates UHHS with potential negative financial incentives (Waters et al., 2015; NQF, 2008) for the pre-existing HACs or HAIs. During the PACT evaluation, a PRE-OP nurse also verifies CSIS inclusion of pre-assessment patient documentation across eight categories: 1) surgery consent; 2) medication reconciliation; 3) medical history / physical exam (H&P); 4) surgical procedure history; 5) patient allergies; 6) problem listing; 7) pre-surgical physician orders; and 8) scheduled procedure / patient education.

Table 2 details the current UHHS recommended pre-assessment KPIs. The five pre-assessment sub-process KPIs are quantitative metrics derived from the CSIS operational data store per the following calculations. % of Patients Seen is the number of PACT patients divided by the number of scheduled out-patient and same-day-surgery cases that should yield 100%. % of Patient Walk-ins is the number of PACT walk-ins divided by the number of PACT visits where the metric should be 5% or less. % of No Shows is the number of no show appointments divided by the total appointment slots and the target is 1% or less. Missing Documentation is a ratio of the total missing documents per category (e.g. 8 total

Table 2. Pre-assessment Sub-process KPIs

Pre-assessment KPIs	BSC	Best Practice
% of Patients Seen	Q	100%
% of Patient Walk-ins	F	5% or less
% of No Shows	F	1% or less
Missing Documentation	Q	0 or none
PACT Length of Stay (LOS)	S	45 min.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

categories) divided by the number of PACT evaluations, where the target is none. *PACT LOS* is the average elapsed minutes for a patient's PACT evaluation and the target is 45 minutes. PACT pre-assessment metrics are analyzed, evaluated, and discussed to minimize future occurrences.

Pre-Operative KPIs

During pre-operative activities, PRE-OP nurses prepare patients for their surgical procedures per physician orders and provide acute patient care after anesthesia. PRE-OP is queue or staging area upstream from intra-operative, with a finite number of areas/beds to prepare patients for surgical procedures. The pre-operative sub-process assists patient's end-state goals of (1) identifying a correct diagnosis for surgical intervention with noted co-morbidities and patient consent, (3) patient exhibits minimal exacerbation of existing disorders, and (4) patient avoids new morbidities. Table 3 details the current UHHS recommended pre-operative KPIs.

The four pre-operative sub-process KPIs are quantitative derivations from the CSIS per the following calculations. *Patient-Arrival-to-Admission* is the patient UHHS arrival time prior to PRE-OP entry and the target timeframe is 2 to 3 hours. *% Patients w/o PACT Eval* is the number patients not seen in PACT prior to PRE-OP, divided by the number of PRE-OP patients, where the target is 5% or less. *PRE-OP Delays* are delay minutes tallied by reason and the average minutes per reason is targeted to be five minutes or less. *PRE-OP LOS* is the elapsed minutes a patient stays in PRE-OP prior to moving to an OR suite and the target average is 35 minutes per patient. All patient experiences of excessive Patient-Arrival-to-Admission time, arriving without a PACT evaluation, PRE-OP delay, or excessive PRE-OP LOS time are analyzed, evaluated, and discussed to avoid future occurrence.

Intra-Operative KPIs

When a patient arrives in an OR suite, an OR nurse begins documenting all people, time, and activities encountered while the patient is present, as required by TJC and CMS, as well as all medication, blood, tissue, and supply usage via the CSIS surgical case. Other CSIS intra-operative documentation includes quality issues for patient longitudinal outcomes, retained object counts, and robotic usage. The final intra-operative CSIS documentation occurs after the patient's surgical case completion when OR staff document the UOS time spent on the OR suite clean up and setup for the next scheduled surgical patient. The intra-operative sub-process assists UHHS in meeting a patient's end-state goals of (2) a patient undergoes the surgical procedure; (3) a patient exhibits minimal exacerbation of existing disorders; and (4) a patient avoids new morbidities.

Table 3. Pre-operative Sub-process KPIs

Pre-operative KPIs	BSC	Best Practice
Patient-Arrival-to-Admission	Q	2 to 3 hours
% Patients w/o PACT Eval	F	5% or less
PRE-OP Delays (By reason)	Q	5 min. or less
PRE-OP Length of Stay (LOS)	S	35 min.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4 details the current UHHS recommended intra-operative KPIs documented via the CSIS during a patient's surgical or endoscopy procedure. The fifteen intra-operative KPIs are derived from the CSIS operational data store per the following calculations. *Prime Time Room Utilization* (e.g. UTIL) is the sum of the time duration to perform each surgical procedure (i.e. "OR In" to "OR Out"), plus the total turnaround time, divided by the prime time available to GENOR and CVOR suites (i.e. 7AM-5PM Mondays to Fridays except 8AM-5PM on Tuesdays) or to HHOR suites (i.e. 7AM-3PM Mondays to Fridays except 8AM-3PM on Tuesdays), with OR availability decreasing on holidays and weekends. The best practice target for Prime Time Room Utilization is 80%. *Modified Block Utilization* is the sum of the time duration to perform each surgical procedure (i.e. preparation of the patient in the OR, anesthesia induction, and emergence) plus the total turnover time, divided by the modified block time scheduled where the target is 75 to 80%. *Block Run-Over Utilization* is the sum of the time duration of surgical procedure performed in 2 hour increments after the assigned prime time block for GENOR and CVOR (i.e., 5PM – 7 PM and 7 PM – 9 PM) or HHOR (i.e., 3 PM – 5 PM and 5 PM – 7 PM) divided by the hours available based on scheduled ORs during those time increments. The Block Run-Over Utilization target is none, with an exception for emergency surgeries. These utilization metrics gauge the use of OR time by specific SSS and surgeons as well as providing the basis for projecting future OR time allotments back to the specific SSS or surgeon.

First Case of the Day On-time Starts (FCOTS) is the percentage of first cases scheduled as of 6 AM day of surgery with an in-room start time (i.e., wheels in) that is either early or not more than 5 minutes after the scheduled start time with a cut off time of 7:30 AM Mondays to Fridays except 8:30 AM on Tuesdays. The target for *FCOTS* is 90% +/- 5 minutes and 95% within 15 minutes. *Subsequent On-time Starts* (e.g. *OTS*) is the percentage of "Prime Time" non-first cases where patient-in-room start time is within 15 minutes of the estimated scheduled procedure start time and the best practice target is 80%

Table 4. Intra-operative Sub-process KPIs

Intra-operative KPIs	BSC	Best Practice
Prime Time Room Utilization	F	80%
Modified Block Utilization	F	75% - 80%
Block Run-Over Utilization	F	None
First Case of the Day On-time Starts	S	90% +/- 5 min. or 95% in 15 min.
Subsequent On-time Starts	S	80% in 15 min.
OR Turnaround Time	F	25-30 min.
In Room to Cut Time	Q	20 – 30 min.
Close to Out of Room Time	Q	12-17 min.
Surgeon RVU Gap	F	30 - 40%
Accurate Case Duration Est.	Q	80%
Day of Surgery Cancel Rate	F	Less than 2%
Day of Surgery Add-ons	Q	Less than 10%
% of Cases with Delays	S	Less than 5%
Average Minutes per Delay	S	5 – 15 min.
Total Minutes Lost to Delays	F	None

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

within 15 minutes. *OR Turnaround Time* (e.g. *TAT*) is the minutes elapsed between the prior patient exiting the room and the succeeding patient entering the room, with the best practice target of 25 to 30 minutes. *FCOTS*, *OTS*, and *TAT* reflect the performance of patient workflow based on the OR schedule and subsequent intra-operative sub-process, communication, and teamwork.

In Room to Cut Time is the elapsed minutes from when a patient enters the room (i.e. wheels in) till the patient's cut-procedure-start time, where the target is 20 to 30 minutes. *Close to Out of Room Time* is the elapsed minutes between the patient's close-procedure-stop time till the patient leaves the OR (i.e. wheels out), where the target is 12 to 17 minutes. *Surgeon RVU Gap* is the minutes elapsed between the prior patient-closed-procedure-stop time and the surgeon's succeeding patient's cut-procedure-start, divided by the prior case time (e.g. cut-procedure-starts to close-procedure-stop). The best practice for Surgeon RVU Gap is 30 to 40%. The first two OR time metrics reflect the total time a patient is in the OR by zones—waiting to start, procedure time, and waiting to exit. Excessive wait times before or after a procedure are reviewed by UHPS administration. Likewise, the Surgeon RVU Gap is a surgical time efficiency measure provided to the surgeon and SSS.

Accurate Case Duration Est. is the percentage of cases where patient-in-room duration (e.g. wheels-in to wheels-out) is within 15 minutes of the estimated in-room duration (i.e. total scheduled surgical time), where the best practice target is 80%. *Day of Surgery Cancel Rate* is a percentage of scheduled cases cancelled after 7PM on the day prior for any reason, divided by the number of scheduled cases for the day of surgery, where the target is less than 2%. *Day of Surgery Add-ons* is a percentage of cases scheduled after 7PM on the day prior, divided by the number of scheduled cases for the day, where the target is less than 10%. These metrics gauge modified block scheduling process (Ryan, et al., 2011) performance and adherence by specific surgeon and SSS.

% of Cases with Delays is a percentage of case delays to cases performed by a surgeon or SSS, where the best practice target is less than 5%. *Average Minutes per Delay* is the elapsed minutes associated with particular SSS delays divided by the number of delays for a SSS, where the target is 5 to 15 minutes. *Total Minutes Lost to Delays* is the number of elapsed minutes associated with delays for a particular SSS, where the best practice target is none. All delayed surgical cases are analyzed, evaluated, and discussed for future delay avoidance.

Post-Operative KPIs

The post-operative sub-process assists in meeting a patient's end-state goals of (3) a patient exhibits minimal exacerbation of existing disorders; (4) a patient avoids new morbidities; and (5) a patient experiences prompt procedure recovery. PACU nurses receive surgical patients from the OR and continue acute care per physician's orders until patient recovery. The PACU has a finite number of acute care beds that accommodate surgical intensive care patients. As surgical patients recover from anesthesia, the need for acute care lessens and patients move to PACU Phase-II ambulatory patient care via the CSIS, leaving acute care patients in acute care beds. Inpatients (IP) recover from anesthesia to be discharged or outpatients (OP) recover to be moved to a hospital bed.

Table 5 details the current UHHS recommended post-operative KPIs documented via the CSIS during a patient's PACU visit. The five post-operative sub-process KPIs are derived per the following calculations. *Bed-Assigned to Ready-to-Move* is for inpatients only and represents the average minutes between the two times, where the target is 15 minutes. *Ready-to-Move to Occupy-Bed* is for inpatients only and represents the average minutes between the two times, where the target is 15 minutes. *Ready-*

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Post-operative Sub-process KPIs

Post-operative KPIs	BSC	Best Practice
Bed-Assigned to Ready-to-Move	Q	15 min. for IP
Ready-to-Move to Occupy-Bed	Q	15 min. for IP
Ready-to-Sign-Out to Sign-Out	Q	5 min. for both
# of Patient Holds in PACU	F	None
PACU Length of Stay (LOS)	S	60 min. – IP or 45 min. - OP

to-Sign-Out to Sign-Out is for both inpatients and outpatients, representing the average minutes between the two times and the target for both is 5 minutes. *# of Patient Holds in PACU* is the number of patient holds in PACU by Patient Type and average elapsed minutes, where the target is none. *PACU LOS* is the average minutes from patient arrival in PACU to discharge from PACU, where the target is 60 minutes for inpatients and 45 minutes for outpatients. PACU KPIs measure how long an inpatient or outpatient spends in PACU, waits to be moved, and denotes patients held in PACU. All excessive times and PACU holds are analyzed, evaluated, and discussed for future resolution.

Central Sterile Supply KPIs

The workflow through CSS moves sterile supplies/instruments/devices to ORs and soiled materials back on all three OR campuses via separate CSS facilities. Prior to a patient's arrival in each OR, CSS delivers a prepared case cart (i.e. up to 8-hours in advance) containing supplies/instruments/devices for the patient's specific surgical procedure per the CSIS SPC pick list and then reprocesses used and un-used case cart contents after the surgical case completion. The CSS sub-process assists the patient's end-state goals of (3) a patient exhibits minimal exacerbation of existing disorders and (4) a patient avoids new morbidities. Likewise, CSS KPIs assist UHHS to avoid HACs and HAIs as well as associated potential negative financial incentives (Waters et al., 2015; NQF, 2008).

Table 6 details the current UHHS recommended KPIs documented during the CSS sub-process. The six CSS KPIs are quantitative and derived per the following calculations. *Immediate Use Steam Sterilization (e.g. IUSS) Usage Rate* is the number of IUSS occurrences divided by the number of surgical procedures, where the target is less than 2%. *Damaged Tray Rate* is the incident total of all trays (e.g., trays of instruments) not meeting surgeon and staff expectations including tracking of wet trays (i.e., sterilized instruments contaminated with retained water), divided by the total trays processed, where the target is none. *Preventive Maintenance (e.g. PM) Plan Adherence Rate* is the percentage of trays receiving the PM review per the contracted number of sets to be reviewed, where the target is 100%. I*nstrument Trays > 25 Lbs.* is the number of tray sets, including vendor trays (e.g., instruments on loan from vendors), exceeding the 25 pound weight limit, where the target is none. *Vendor Tray Management Non-Compliance (e.g. NC) Rate* is the number of incidents that the vendor tray management process was not met, where the target is none. *Surgeon Preference Card (e.g. SPC) Update Review* is the number of surgical preference cards signed off by the SSS surgeons every 6 months divided by the number of SSS SPCs, where the target is 100%. These operational and tactical metrics impact strategic CSFs of patient safety, patient quality of care, employee satisfaction, and hospital margin.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. CSS Sub-process KPIs

Central Sterile Supply KPIs	BSC	Best Practice
IUSS Usage Rate	Q	Less than 2%
Damaged Tray Rate	Q	None
PM Plan Adherence Rate	Q	100%
Instrument Trays > 25 Lbs.	F	None
Vendor Tray Mgt. NC Rate	S	None
SPC Update Review	F	100%

KPI DATA VISUALIZATION

Figures 1, 2, and 3 are examples of data visualization for pre-assessment, intra-operative, and CSS sub-process KPIs. KPI filters are depicted on the left side of each figure. Figure 1 takes each of the 8 KPI categories of PACT *Missing Documents* required for complete documentation and uses line graphs to illustrate the magnitude of the top three missing document frequency counts above 30% for medical reconciliation, problem lists, and surgical procedure history. Figure 1 also shows the bottom five categories that were 15% or less. Figure 2 takes the *Prime Time Utilization* KPI and displays it as a stacked bar graph and a pie chart, along with a bar chart of modified block minutes assigned to each SSS. Lastly, Figure 3 displays *IUSS* KPIs via line graphs and *Tray Set* KPIs via bar charts. All of the perioperative sub-process KPI data residing in the CSIS data mart can be visualized like these examples.

Intra-Operative Performance Dashboard

Before FY2016, UHPS tallied perioperative BSC measures into electronic dashboards and pushed the results out to stakeholders (Ryan et al., 2014). As of FY2016, KPI data visualizations and dashboards are pulled on-demand by stakeholders. Figure 4 below depicts a dashboard query of intra-operative metrics. For a given time period, location, OR room, or SSS, the dashboard query will visualize intra-operative KPIs of *FCOTS* and *OTS*. Complete cases have been UHPS verified and the *Add-on Case* gauge reflects the criticality of 25% add-on case volume to resource capacities. *Case Cancelled on DOS* is a Pareto chart identifying surgeons who have the most cases cancelled on days which specific surgeries were scheduled to occur. The *Patients entering PACU* graph depicts cumulative peak volume times of patients entering PACU over a 24-hour frequency. Lastly, the *Number of Cases with Late Starts* is a Pareto chart identifying surgeons with the highest frequency of surgical cases that did not start on-time.

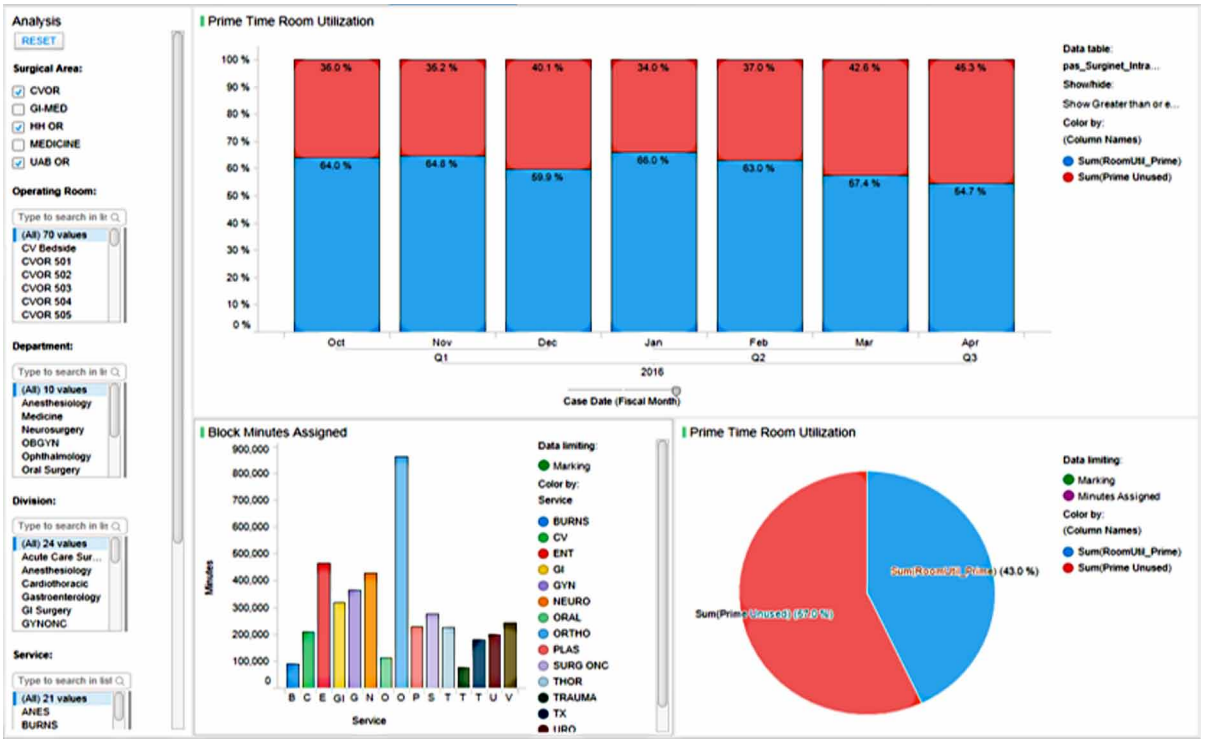
SUMMARY AND DISCUSSION

Perioperative stakeholders pull KPI data visualizations and dashboards on demand as needed. The 35 perioperative KPIs represent distinct perioperative sub-process perspectives from which stakeholders can gauge, visualize, and measure performance while providing the opportunity to drill down to potential problem root causes within a given sub-process rather than being limited to visibility of associated symptoms elsewhere in the process. In regards to a BSC approach (Kaplan & Norton, 1996), 13 of the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. PACT Missing Document KPIs

Figure 2. Prime Time Utilization KPI



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Central Sterile Supply KPIs

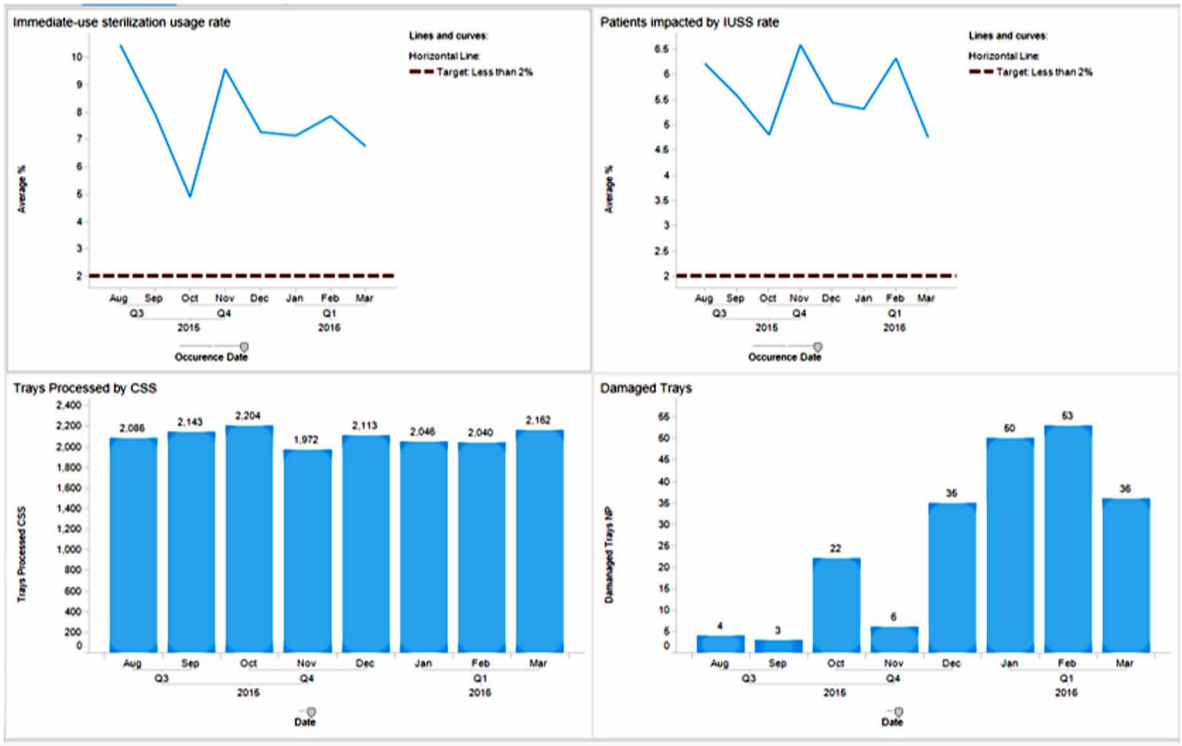
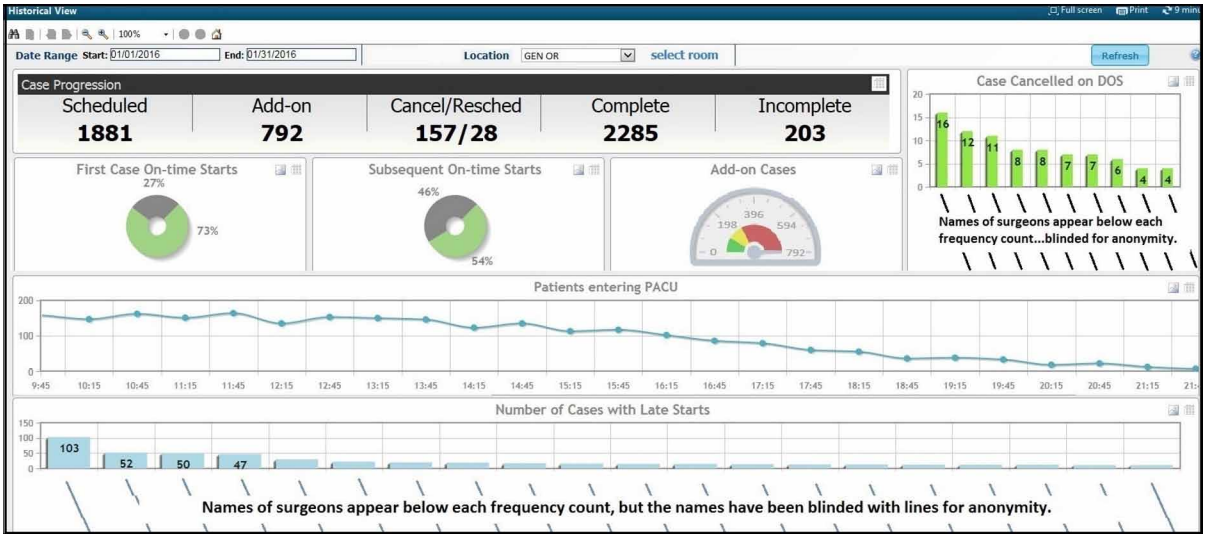


Figure 4. Intra-operative Performance Dashboard



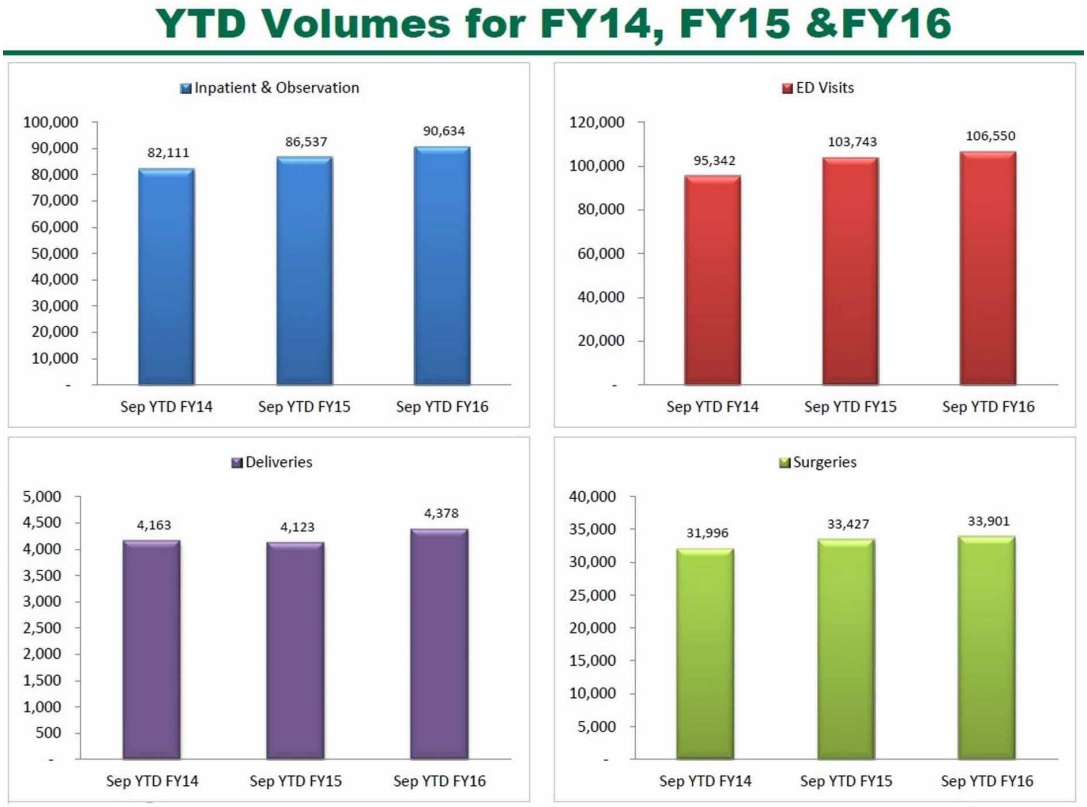
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

KPIs reflect financial perspectives (e.g., 37.1%), 14 reflect quality of the process (e.g., 40%), and 8 reflect satisfaction of the customer (e.g., 22.9%). Similarly, all 35 KPIs provide a learning perspective to evaluate each perioperative sub-process as well as identify how each KPIs in each sub-process support specific patient end-state goals. Furthermore, breaking down perceived perioperative complexity (Fowler et al., 2008) into component sub-processes demonstrates a proven approach to understanding complexity by reducing the perspective of the phenomenon into smaller more manageable, comprehendible, and explainable units (e.g., sub-processes).

The BPM taskforce's charge to create actionable information to support initiatives at the operational, tactical, and strategic levels requires a periodic iterative review of KPIs due to anticipated improved understanding of specific sub-process relationships. Improved sub-process understanding will drive the need to revise existing KPIs to better reflect the improved comprehension. Hence, periodic KPI reviews and revisions provide continuous improvement opportunities for each sub-process and the overall perioperative process workflow.

To demonstrate increased perioperative patient flow, the GENOR (i.e., North Pavilion campus without CVOR) surgical volume increased by 11.69% between FY2006 (18,921 cases) to FY2014 (21,133 cases). Figure 5 reflects UHHS increased patient volume over FY2014, FY2015, and FY2016, with increasing patient flow across inpatient and observation (10.37%), emergency department (11.76%), surgical procedures (5.95%) as well as labor and delivery (5.16%).

Figure 5. Patient Volumes FY2014, FY 2015, and FY 2016



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

This study highlights balanced scorecard metrics from each perioperative sub-process (e.g. pre-assessment, pre-operative, intra-operative, post-operative, and ancillary CSS activities), demonstrates how the metrics are applicable as KPIs, and defines how each perioperative sub-process supports specific patient end-state goals. Empowered individuals, integrated IS, and a holistic framework for analysis, evaluation, and synthesis of sub-process measures prescribed an a priori environment for replication and use. Moreover, BPM practices were adaptable to explain overall perioperative complexity and improvement efforts within the perioperative sub-processes. The cycle of analysis, evaluation, and synthesis also reinforced communication and stimulated individual as well as collective organizational learning.

Our case study contributes to the healthcare IT literature by examining how continuous improvement is applicable to BPM practices via establishing KPIs as well as the management of perioperative sub-processes nested within the hospital environment. This paper also fills a gap in the literature by identifying sub-process KPIs to help explain perioperative complexity and how process metrics as data are both a performance measure and management tool.

This study was limited to a single case, where future research should broaden the focus to address this issue along with others that the authors may have inadvertently overlooked. The case examples presented in this study can serve as momentum for BPM via KPI identification and use in healthcare methodology, comprehension, and extension. The study's results should be viewed as exploratory and in need of further confirmation. Researchers may choose to further or expand the investigation, while practitioners may apply the findings to create their own version of BPM through KPI identification and use within the hospital environment.

REFERENCES

Ackoff, R. L. (1967). Management misinformation systems. *Management Science*, *14*(4), 147–156. doi:10.1287/mnsc.14.4.B147

Albanese, M. P., Evans, D. A., Schantz, C. A., Bowen, M., Disbot, M., Moffa, J. S., ... Polomano, R. C. (2010). Engaging clinical nurses in quality and performance improvement activities. *Nursing Administration Quarterly*, *34*(3), 226–245. doi:10.1097/NAQ.0b013e3181e702ca PMID:20562572

Baker, J. (2015). Language of Improvement: Metrics, Key Performance Indicators, Benchmarks, Analytics, Scorecards, and Dashboards. *AORN Journal*, *102*(3), 223–227. doi:10.1016/j.aorn.2015.07.015 PMID:26323219

Blumenthal, D. (2012). Performance improvement in healthcare—seizing the moment. *The New England Journal of Medicine*, *366*(21), 1953–1955. doi:10.1056/NEJMp1203427 PMID:22533534

Catalano, K., & Fickenscher, K. (2007). Emerging technologies in the OR and their effect on perioperative professionals. *AORN Journal*, *86*(6), 958–969. doi:10.1016/j.aorn.2007.07.007 PMID:18068401

CMS. (2015). Hospital-Acquired Conditions List 2015. Retrieved from http://www.cms.gov/medicare/medicare-fee-for-service-payment/hospitalacqcond/hospital-acquired_conditions.html

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Doebbeling, B. N., Burton, M. M., Wiebke, E. A., Miller, S., Baxter, L., Miller, D., ... & Pekny, J. (2012). Optimizing perioperative decision making: improved information for clinical workflow planning. *Proceedings AMIA Annual Symposium* (pp. 154–163).

Dougherty, D., & Conway, P. (2008). The 3Ts road map to transform US health care. *Journal of the American Medical Association*, *299*(19), 2319–2321. doi:10.1001/jama.299.19.2319 PMID:18492974

Duro, M. (2015). In focus: Factors that affect surgical instrument availability and turnover time. *AORN Journal*, *101*(1), C7–C8. doi:10.1016/S0001-2092(14)01326-X PMID:25689870

Eckerson, W. (2006). Deploying dashboards and scorecards - TDWI best practice report. *The Data Warehousing Institute*. Retrieved from http://www.microstrategy.com/Strategy/media/downloads/solutions/TDWI-Best-Practices-Report-Deploying-Dashboards-and-Scorecards.pdf

Waters, T. M., Daniels, M. J., Bazzoli, G. J., Perencevich, E., Dunton, N., Staggs, V. S., ... & Shorr, R. I. (2015). Effect of Medicare's nonpayment for hospital-acquired conditions: Lessons for future policy. *Journal of the American Medical Association Internal Medicine*, *175*(3), 347–354. PMID:25559166

Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Fowler, P., Craig, J., Fredendall, L., & Damali, U. (2008). Perioperative workflow: Barriers to efficiency, risks, and satisfaction. *AORN Journal*, *87*(1), 187–208. doi:10.1016/j.aorn.2007.07.001 PMID:18184599

Jeston, J., & Nelis, J. (2008). *Business Process Management: Practical Guidelines to Successful Implementations* (2nd ed.). Burlington, MA: Elsevier, Ltd.

Jones S., Rudin R., Perry T., & Shekelle P. (2014). Health information technology: an updated systematic review with a focus on meaningful use. *Annuals of Internal Medicine, 160*(1), 48-54.

Kanich, D. G., & Byrd, J. R. (1996). How to increase efficiency in the operating room. *The Surgical Clinics of North America*, *76*(1), 161–173. doi:10.1016/S0039-6109(05)70429-1 PMID:8629198

Kaplan, R., & Norton, D. (1996). *The Balanced Scorecard: Translating Strategy into Action*. Boston: Harvard Business School Press.

Karami, M., Safdari, R., & Rahimi, A. (2013). Effective radiology dashboards: Key research findings. *Radiology Management*, *35*(2), 42–45. PMID:23638580

Karimi, J. (1988). Strategic planning for information systems: Requirements and information engineering methods. *Journal of Management Information Systems*, *4*(4), 5–24. doi:10.1080/07421222.1988.11517805

Macario, A. (2006). Are your hospital operating rooms efficient? A scoring system with eight performance indicators. *Anesthesiology*, *105*(2), 237–240. doi:10.1097/00000542-200608000-00004 PMID:16871055

Macario, A., Vitez, T., Dunn, B., & McDonald, T. (1995). Analysis of hospital costs and charges for inpatient surgical care. *Anesthesiology*, *6*(83), 1138–1144. doi:10.1097/00000542-199512000-00002 PMID:8533904

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Marjamaa, R., Vakkuri, A., & Kirvela, O. (2008). Operating room management: Why, how and by whom? *Acta Anaesthesiologica Scandinavica*, *52*(5), 596–600. doi:10.1111/j.1399-6576.2008.01618.x PMID:18419711

Munroe, M. C., & Wheeler, B. R. (1980). Planning, critical success factors, and managements information requirements. *Management Information Systems Quarterly*, *4*(4), 27–37. doi:10.2307/248958

National Quality Forum (NQF). (2008). Serious reportable events. Retrieved from http://www.quality-forum.org/Publications/2008/10/Serious_Reportable_Events.aspx

Paré, G. (2001). *Using a positivist case study methodology to build and test theories in information systems*. Montréal, Canada: HEC Montréal Business School. Retrieved from http://expertise.hec.ca/gresi/wp-content/ uploads/2013/02/cahier0109.pdf

Peters, J., & Blasco, T. (2004). Enhancing hospital performance through perioperative.

PwC Health Research Institute. (2012). The future of the academic medical center: strategies to avoid a meltdown. Retrieved from http://www.pwc.com/us/en/health-industries/publications/the-future-of-academic-medical-centers.jhtml

Rockart, J. (1979). Chief executives define their own data needs. *Harvard Business Review*, *57*(2), 81–93. PMID:10297607

Ryan, J., Doster, B., Daily, S., & Heslin, M. (2008). Soft innovation as data-driven process improvement exploited via integrated hospital information systems. *Proceedings of the 41st Hawaii International Conference on System Sciences*. IEEE Computer Society Press. 10.1109/HICSS.2008.405

Ryan, J., Doster, B., Daily, S., & Lewis, C. (2011). Analyzing block scheduling heuristics for perioperative scheduling flexibility: a case study perspective. *Proceedings of the 44th Hawaii International Conference on System Sciences*. IEEE Computer Society Press. 10.1109/HICSS.2011.64

Ryan, J., Doster, B., Daily, S., & Lewis, C. (2012). Evaluating and improving the perioperative process: benchmarking and redesign of preoperative patient evaluations. *Proceedings of the 45th Annual Hawaii International Conference on System Sciences*. IEEE Computer Society Press. 10.1109/HICSS.2012.250

Ryan, J., Doster, B., Daily, S., & Lewis, C. (2014). A balanced perspective to perioperative process management aligned to hospital strategy. *International Journal of Healthcare Information Systems and Informatics*, *9*(4), 1–19. doi:10.4018/ijhisi.2014100101

Ryan, J., Doster, B., Daily, S., & Lewis, C. (2016). A case study perspective for balanced perioperative workflow achievement through data-driven process improvement. *International Journal of Healthcare Information Systems and Informatics*, *11*(3), 19–41. doi:10.4018/IJHISI.2016070102

Silverman, D., & Rosenbaum, S. (2009). Integrated Assessment and Consultation for the Preoperative Patient. *The Medical Clinics of North America*, *93*(5), 963–977. doi:10.1016/j.mcna.2009.05.010 PMID:19665614

Tarantino, D. (2003). Process Redesign Part 1: Process Selection. *Physician Executive*, *29*(6), 71–73. PMID:14686251

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tenner, A., & DeToro, I. (1997). *Process redesign: the implementation guide for managers*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Turban, E., Sharda, R., Aronson, J., & King, D. (2008). *Business Intelligence: A managerial approach*. Upper Saddle River, New Jersey: Prentice Hall.

Walton, M. (1986). *The Deming Management Method*. New York: Dodd, Mead.

Wears, R., & Berg, M. (2005). Computer technology and clinical work: Still waiting for godot. *Journal of the American Medical Association*, *293*(10), 1261–1263. doi:10.1001/jama.293.10.1261 PMID:15755949

Weber, R. (2004). The rhetoric of positivism versus interpretivism: A personal view. *Management Information Systems Quarterly*, *28*(1), iii–xii.

Wilbanks, B., & Langford, P. (2014). A review of dashboards for data analytics in nursing. *CIN: Computers, Informatics, Nursing*, *32*(11), 545–549. PMID:25310224

Wright, J., Roche, A., & Khoury, A. (2010). Improving on-time surgical starts in an operating room. *Canadian Journal of Surgery*, *53*(3), 167–170. PMID:20507788

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, California: Sage Publications.

Zani, W. M. (1970). Blueprint for MIS. *Harvard Business Review*, *48*(6), 85–90.

This research was previously published in the International Journal of Healthcare Information Systems and Informatics (IJHISI), 12(4); edited by Joseph Tan and Qiang (Shawn) Cheng; pages 13-30; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 88
Building Intelligent Systems for Paying Healthcare Providers and Using Social Media to Detect Fraudulent Claims

Jack S. Cook
The College at Brockport, USA

M. Pamela Neely
The College at Brockport, USA

ABSTRACT

Using an interpretive case study approach, this paper describes the data quality problems in a regional health insurance (RHI) company. Within this company, two interpretive cases examine different processes of the healthcare supply chain and their integration with a business intelligence system. Specifically, the first case examines RHI's provider enrollment and credentialing process, and the second case examines the processes within the special investigations unit (SIU) for investigating and detecting fraud. The second case examines DIQ issues and how social media can be used to acquire evidence to support a fraud case. In addition, the second case utilized lean six sigma to streamline internal processes. A data and information quality (DIQ) assessment of these processes demonstrates how a framework, referred to as PGOT, can identify improvement opportunities within any information intensive environment. This paper provides recommendations for DIQ and social media best practices, and illustrates these best practices within this real-world context of healthcare.

INTRODUCTION

Organizations rely on data for a multitude of applications: customer service and relationship management, decision making, and regulatory compliance. There is currently a push specifically in healthcare to be smarter about its management of key performance outcomes, as well as to actively investigate fraud

DOI: 10.4018/978-1-7998-1204-3.ch088

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and potential fraud. Organizations are beginning to search for methods to deliver information to decision makers in more intelligent ways. A desire to lower costs and improve patient outcomes is driving this need for more intelligent data. Therefore, one option is to use business intelligence (BI) and/or collective intelligence (CI) to form the heart of any system that delivers organizationally sustaining data to providers and payers so they can make decisions that positively impact patient and service outcomes, and lower costs.

The challenge is to convince administrators, clerical staff, clinicians and physicians to buy-in to the importance of integrating previously disparate data to create a real-time view of the relationship between a patient, the provider, and the payer or payers. This article focuses on BI and CI systems that support healthcare, and the data and information quality (DIQ) issues that are inherent in any environment that is information intensive. These systems have tremendous potential to impact healthcare, from multiple perspectives – patients, providers and payers. From a patient's perspective, BI can provide clinicians treatment and immunization recommendations that are preventative in nature, improving both diagnostic accuracy and better implemented care plans. Billing is critical in today's healthcare environment; providers transmit claims to multiple payers, and payment denials are common. BI dramatically improves data and information quality. Therefore, from the provider's perspective, IT professionals can streamline the complex billing process using BI. From a payer's perspective, a better connection to providers and a more intelligent way to enroll and credential them will reduce claims adjustments and disputes. CI has the ability to systematize the discovery of fraud – both provider fraud and member fraud.

This article provides a brief literature review of healthcare, business intelligence (BI), and collective intelligence (CI), as well as data and information quality (DIQ), followed by a brief summary of fraud prevention, deterrence, detection and response. The article then examines two cases that make the point that BI and CI are reasonable approaches to managing key healthcare performance outcomes, including early detection of fraud. The first case details the process that a payer performs to enroll and credential a healthcare provider. The second case examines a special investigations unit's process for determining if fraud has occurred, independent of whether it was committed by a provider or member. The objective in this second case was to streamline the process by utilizing lean six sigma to eliminate waste and explore fraud red flags using social media so that more investigations, using the same resources, can be processed within a given time period. We conclude with recommendations for using BI and CI most effectively in a healthcare setting.

LITERATURE REVIEW

Healthcare

As defined by the U.S. Department of Health and Human Services, a healthcare provider is "a provider of services as defined in §1861(u) of the Act (Social Security Act), a provider of medical or health services as defined in §1861(s) of the Act, and any other person or organization who furnishes, bills, or is paid for healthcare services or supplies in the normal course of business." (U.S. Department of Health and Human Services, 2001) Until recently, healthcare providers were paid when sick people sought treatment. Quality and outcomes were not rewarded, but rather providers were paid based on how much was done to treat the patient. In the future, healthcare providers will be given financial incentives to keep people well, and quality, not treatment, will matter a great deal. This shift from volume to value will require a rethinking of treatment plans and a shift towards preventative care.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Business Intelligence (BI)

Traditionally, BI has been described as supporting human intelligence with "technologies, applications and processes for gathering, storing, accessing and analyzing data to help users make better decisions" (Wixom & Watson, 2010, p. 14). BI is used in many industries, including healthcare (Olinsky & Schumacher, 2010; Sillup, Klimberg, & McSweeney, 2010).

Wixom and Watson (2010) contend that BI is about getting the data in (via a data warehouse or data mart) and getting the information out (using organizational specific tools to meet decision making purposes). They provide a model of a BI-based organization that includes the traditional data warehouse elements and data integration coupled with analytical tools that have historically been of the online analytical processing (OLAP) and decision support type. These historical tools are now under the umbrella of BI technologies and applications, and are not listed separately. Additionally, and equally as important, the model includes processes related to metadata, data quality and governance. The recognition of these processes as important to the BI environment reinforces the need for practitioners to evaluate BI as a system, not simply a technology.

An important aspect of BI is the ability to analyze data and return information that can be used in decision making. Related to BI is the field of big data analytics. As data has become more abundant, there is a need for new technologies and techniques to analyze this data effectively (Chen, Chiang, & Storey, 2012). Another challenge in healthcare is the abundance of free text that is not easily analyzed. Doctors' notes, prescription information, and notes attached to test results can all be found in the typical healthcare system. New technologies in text analytics are necessary to adequately mine the data to affect decision making (Chen et al., 2012).

Chang (2014) introduces the concept of Business Intelligence as a Service (BIaaS) in the Cloud. He implements the approach using two financial based scenarios; risk and pricing for financial analysis and stock market analysis. This approach harnesses the power of the Cloud and allows for more complex models and improved accuracy of predictions. Taken together, the structure of BI is morphing and becoming an even more powerful tool.

Social Media

This term refers to much more than just Facebook. The wealth of user-created information on the web (Web 2.0) has led to the ability to share virtually anything with anyone. The various categories of social media enable end users to interact with each other in the virtual worlds, sharing opinions, discussions, pictures, videos, and factual information. The various parties may or may not know each other in the face-to-face world. Thus, social media has broadened the opportunities to meet and collaborate with people from around the world. It has also opened up the possibilities of employers learning more about employees than has previously been possible. However, the large amount of data generated by social media will need new tools in order to truly harness the power of all of this data. Chang (2016) describes a Social Cloud that allows companies to learn how people react to products and services. Because the tool permits the use of multiple APIs, this is a good first step in designing systems that will capture data that may be related to fraud.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Collective Intelligence (CI)

CI is the shared or group intelligence that emerges from consensus decision making. These group decision outcomes depend on rules, such as majority rule or quorum rule (Hastie and Kameda, 2005). Multiple medical/healthcare related studies use collective intelligence to show that group decisions are consistently better than individual decisions (Wolf, et al, 2015; Engel, 2014; Henandez-Chan et al, 2012; Jane Budge et al, 2015, Kurvers et al, 2015). Crowdsourcing and social media support CI, and the collective behavior of a group of individuals online in terms of their social media usage can provide evidence of fraudulent behavior on the part of one or more members in the group.

Data Quality Management (DQM)

Weber, Otto and Osterle, (2009) define DQM as quality-oriented management of data as an asset, that is, the planning, provisioning, organization, usage, and disposal of data that supports both decision-making and operational business processes, as well as the design of the appropriate context, with the aim to improve data quality on a sustained basis. The management of DIQ is also referred to as data governance.

The quality of data and information is dependent on how it is used, also known as fitness-for-use (Neely & Cook, 2011). It is difficult to trust your business intelligence system when DIQ is poor, and processes are ill conceived. Within healthcare, a reduced number of denials for claim payment and increased patient care (Leape et al., 1993) can result from improved DIQ because the system can notify providers of immunization and care plans automatically. This should result in better patient satisfaction. Fewer adjustments to the claims processing system (Romano Mj, 2011), decreased time to enroll providers, and reduced labor with respect to enrolling providers will result in better provider satisfaction and an increase in revenues in general (Rauscher & Wheeler, 2008).

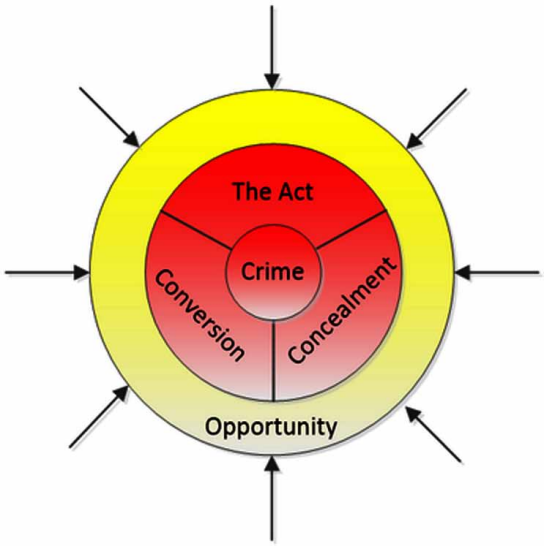
DIQ in healthcare will continue to be of the utmost importance, for patient safety and satisfaction as well as the financial health of the system. Given the unique nature of the payment system in healthcare, it is critical that systems facilitate providers being paid. Getting paid from sometimes multiple payers, when payers have the ability to deny payment for a multitude of reasons, is much more challenging than in almost any other industry.

Fraud Management

Fraud prevention, deterrence, detection and response has been extensively studied in the literature (Bonchi et al, 1999; Debreceny et al, 2005; Fawcett & Provost, 1997). In particular, fraud is prevalent in healthcare (Bosler, 2015; Bourquard & Kirsch, 2015; Delancy & Gorton, 2015; Simborg, 2014). As seen in Figure 1, fraud consists of the criminal act, conversion, and concealment, assuming the opportunity to commit the crime exists in the first place. In Figure 1, the larger circle represents an opportunity for someone to commit a fraud. The larger the opportunity, the more likely it is that someone will commit a crime. If no opportunity exists, then no crime can be committed. The arrows pointing in towards the opportunity circle represent the actions taken by the company to close the "hole", represented by increased security measures, surveillance, and utilizing social media to name a few. Given the opportunity, the fraudster will commit the criminal act, conceal it from the organization, and eventually convert it to a form (primarily money) where it can be used by the fraudster.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. The crime



Dorminey et al (2008) explores the topic from the perspective of the fraud triangle as well as a meta-model of white collar crime. When the fraud triangle was developed, it was seen as an employee/employer relationship and occurred within the context of the work environment. Often these people were crisis fraudsters, and this was the type that was originally envisioned when the original fraud triangle was proposed. This person has pressure, rationalizes their action and has the opportunity to commit the fraud.

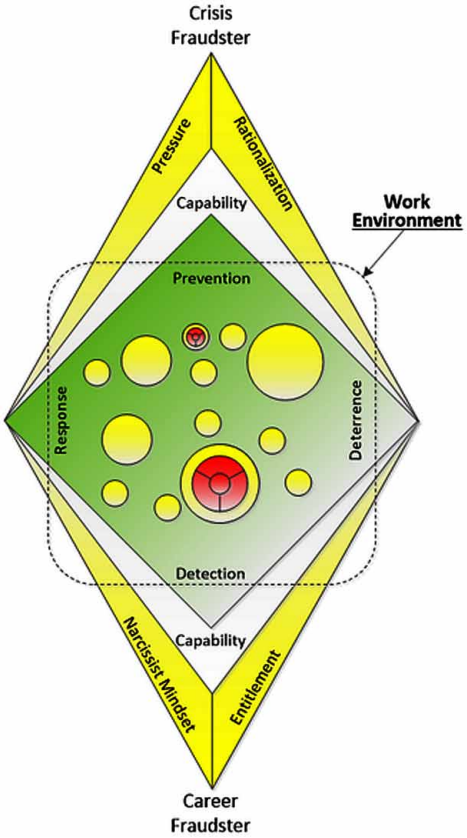
We have extended this work by identifying two different types of fraudsters, with differing motivations – crisis fraudster and the career fraudster. The fraud compass (as seen in Figure 2) shows that both types of fraudsters need to possess the capabilities necessary to commit the criminal act. Also, our model shows the relationship between the crime (fraud) and the work environment. As online resources, especially social media, have become more prevalent, the work environment has expanded from the physical space to the virtual space. As one work environment becomes too challenging to defraud, the career fraudster moves to another company, committing frauds in each position they hold. They have a narcissistic mindset and a feeling of entitlement, which is very different from what was envisioned by the fraud triangle. Until they are prosecuted and convicted, they will continue to seek opportunities to commit fraud. In order for fraud to occur, both types of fraudsters must have the capability to commit the fraud and an opportunity. Luckily, companies can limit the opportunities through better use of BI, CI, and social media to identify fraudulent acts earlier.

Interplay between DIQ, CI and BI

Quality can be dramatically improved; for example, medical errors account for the deaths of at least 200,000 people in the U.S. even though the US healthcare system has traditionally been known for high quality (Harmon, 2009). BI systems could help to alleviate many of the errors found in healthcare systems. For example, diagnostic errors could be mitigated with BI by utilizing data mining techniques to look for cause and effect of lifestyle factors. These lifestyle factors could guide the diagnostician towards additional testing, ultimately resulting in a correct diagnosis.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Fraud compass



By using decision-support tools or BI, these systems can increase the adherence to protocol-based systems. However, these systems can be costly and there is a need to study the return on investment. Bardhan and Thouin (2012) show that the use of electronic health record systems (EHRs) can positively impact the treatment for heart attacks, heart failures and pneumonia. Furthermore, using financial management systems can result in lower operating expenses. On the other hand, Romano (2011) studied clinical decision making systems (CDS) and did not find a correlation between visits with and without a CDS on the quality of care. Linder et al. (2007) had a similar finding. Their study examined EHR use and ambulatory care quality and found no significant association. According to Giannangelo and Fenton (2008), in order for the EHRs to be effective, there must be a standard of codification and it must be used consistently to ensure that data is correct and reliable. This codification, particularly within the financial portion of the system, can lead to more effective revenue cycle management (RCM).

RCM is the collection of methods used to manage claims payment processing and revenue generation. Unlike the standard business-to-business or business-to-consumer model, whereby an invoice or request for payment is generated and the payment is made, a significant proportion of the revenue in healthcare is derived from non-patient payment (e.g., a third party), resulting in multiple streams of payments that must be managed. RCM is a frequent component of BI, allowing managers to increase the speed and amount of patient revenue collection (Rauscher & Wheeler, 2008).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

There is potential for social media and crowdsourcing to support collective intelligence, thus allowing early detection of fraud. The BI system can automatically scan social media sites, allowing auditors additional support for decision making, effectively providing a larger sample of data and providing the majority rule components of collective intelligence (Bardhan and Thouin 2012).

A properly implemented business intelligence system can play a vital role in ensuring quality throughout the data and information lifecycle. Thus, this article examines the connections between these three very important and mutually dependent areas of practice, BI, CI, and DIQ, in such a data-rich environment as healthcare. The PGOT framework, discussed next, provides an excellent mechanism for evaluating opportunities for improving quality.

DATA AND INFORMATION QUALITY FRAMEWORK – PGOT

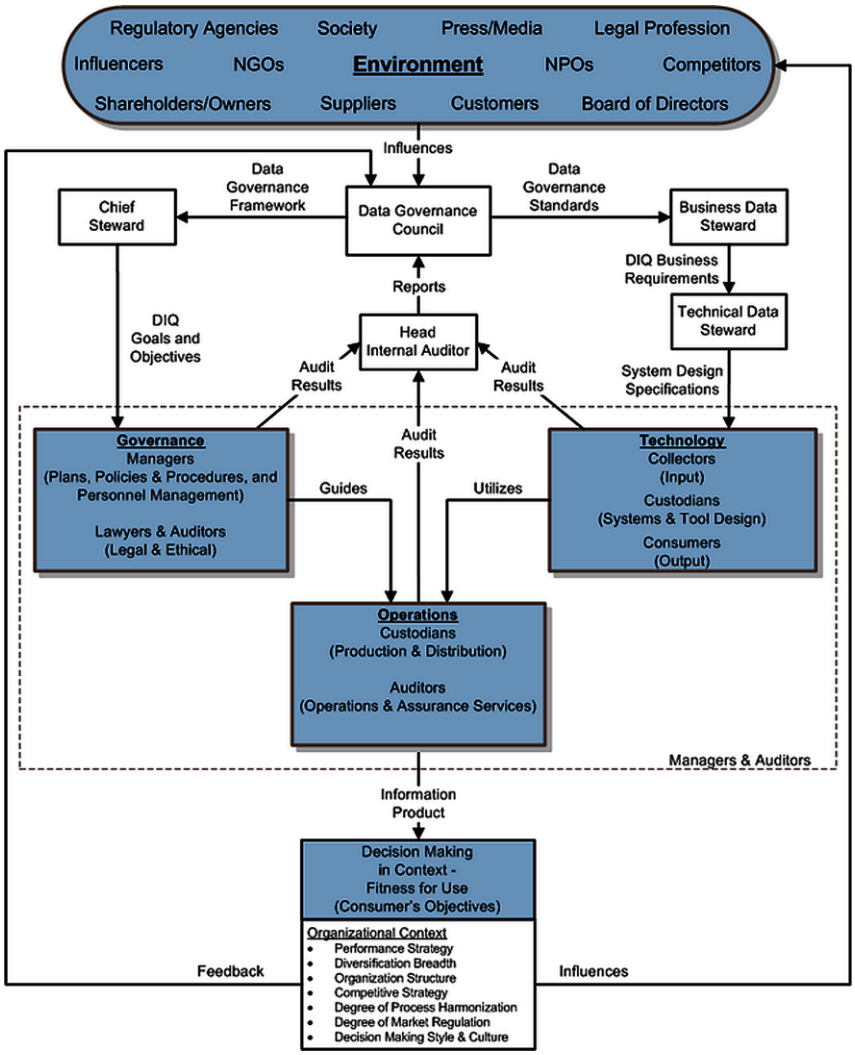
As evidenced by frameworks such as COBIT (Boritz, 2005), IT governance has been at the forefront of management concerns for a number of years. Increasingly, the challenge is to govern not only the systems, but also the data residing in the systems (Maguire, 2008). As indicated earlier, DIQ is dependent on how the data is used. Thus, acceptable quality in one instance may be unacceptable in another. For example, prescription information may be entered into a system for a patient. The data is accurate, but incomplete. The system has not captured the fact that the patient is seeing a specialist who has prescribed other medications that negatively interact with what the primary care physician has prescribed. Proper data governance, across all care providers, could help to alleviate this problem.

Compared with the extensive research literature on IT governance, academic interest and research on data governance and its relationship to business intelligence is still in its infancy. As introduced in Neely and Cook (2011), the PGOT (***p***eople, ***g***overnance, ***o***perations, and ***t***echnology) framework provides a foundation for evaluating the quality of data and information. Figure 3 presents an enhanced version of the PGOT with an increased emphasis on the people within the system. The governance plan, operational efficiency and technical competency of the systems continue to be a driving force in the overall quality of the data and information. Based on the contingency theory of organizational design, the underlying assumption of the PGOT framework is that there is no universal design applicable to all organizations for data and information quality, and healthcare is no exception. It is the contextual factors in healthcare that impact the contribution of data governance in enhancing IT governance and organizational performance. This framework focuses on long-term data and information quality improvement and error prevention, rather than correction. In the two cases discussed later, business intelligence systems are the mechanism by which DIQ is improved. The usefulness of a BI system is directly correlated to the quality of the data that it contains and the information it produces.

It is impossible to ensure "perfect" data quality on all dimensions given that the quality of data and information is dependent on its use; that is, decision making in context. The role of the data governance council, and by extension the chief data steward, business data steward and technical data steward, is to ensure the data and information is of sufficient quality to best meet the demands of most decision makers. The business data steward must define the business requirements and expectations, and trust the technical data steward to work with developers to build a system that meets data and information quality-related business requirements. In addition, the chief data steward must design and implement plans, policies and procedures that comply with legal and ethical requirements and meets the needs of data consumers. The chief data steward would also design and implement plans for using social media to detect potential fraud.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. DIQ in context: PGOT research framework



Although perfect data is impossible, there are multiple feedback loops illustrated throughout the framework, many of them resulting from data quality audits, to help in the identification of opportunities to continuously improve DIQ and detect fraud. Therefore, the ability to audit the data and information quality from an operational, technical, and compliance perspective is vital to DQM. As a result, the BI system should have embedded continuous auditing controls, and metrics from such audits should be frequently analyzed for system usability and reliability.

The roles of the people who influence, manage and use the BI system with an emphasis on the data governance system will be illustrated later in the second case. The PGOT framework will be used to focus attention on DIQ within the BI system, specifically when investigating fraud. The next section briefly describes the healthcare environment and this is followed by a description of two case studies, illustrating how a methodical approach to examining DIQ can help to improve the quality and usefulness of the BI system, and ultimately positively impact healthcare financials.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BETTER MANAGING COSTS AND MEDICAL OUTCOMES USING DIQ, CI AND BI

Medicare and Medicaid continue to ratchet down provider payments, and this is not expected to change. Fee for service is becoming less prevalent, while value-added and bundled services, referred to within the industry as episodic payments, will increase. Many payers already pay a lump sum for maternity. Physician procedural income will decline while their employment and integration with healthcare facilities will increase. As a result, hospitals, physicians and other healthcare providers will either find a way to deliver healthcare for less money or go out of business. The fundamental challenge of healthcare reform is to simultaneously improve the quality of care while lowering its costs. Although profitability is a dirty word in healthcare, without it, healthcare will not have the funds needed to provide better care for patients, whether it is acquiring the latest technology or hiring adequate staff to treat patients.

Case 1: Provider Relations

This case involves a regional health insurance company and a large, independent provider group. This regional health insurance company, referred to here as RHI, is located in the United States and is a nonprofit health insurer. It has provided health insurance for over eight decades and currently serves approximately two million members. The provider group represents almost 1,000 specialists and primary care providers. It assists in negotiating payment rates with third parties, and facilitates credentialing its providers. The specialist and primary care provider (SPCP) group will be referred to here as the SPCP group for short.

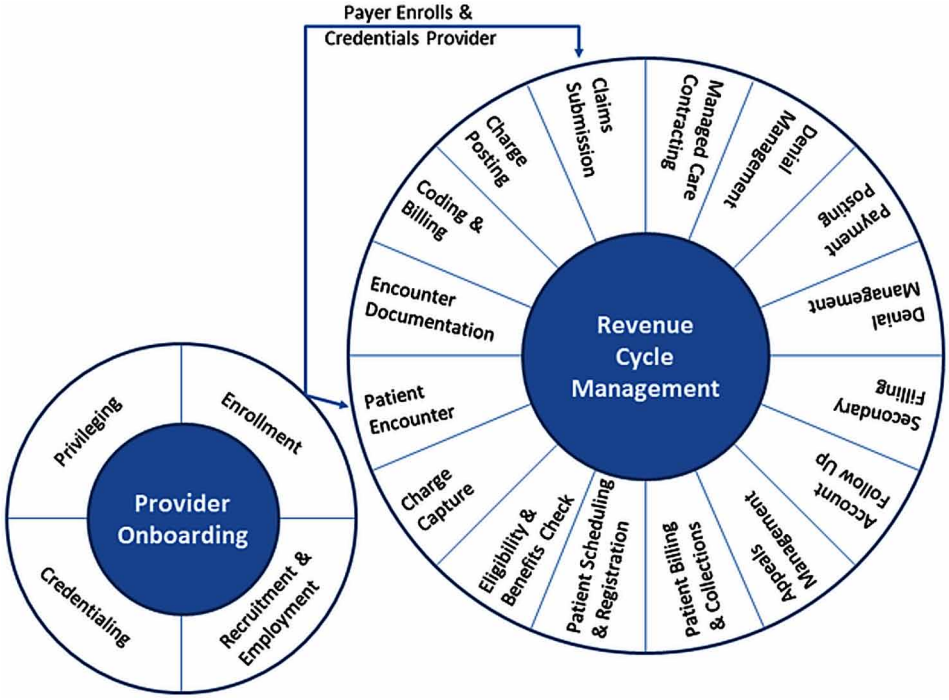
As seen in Figure 4, there is a relationship between the provider's onboarding process (i.e., the process whereby a physician or specialist is authorized by the organization to treat patients) and revenue cycle management. This relationship ultimately dictates when a provider will be paid for services rendered. A provider must first be credentialed by both the healthcare organization and the payer to ensure that that they have the necessary qualifications to practice medicine. Privileging is the process of determining what services a provider may offer patients. Once credentialing and privileging are complete, the provider is enrolled. From a payer's perspective, a claim that is submitted must come from an enrolled provider. Otherwise, the claim is denied, and if the provider should have indeed been paid and was not due to the payer's inefficient credentialing process, additional work will need to be completed to retroactively pay the claim. From both the provider's and payer's perspective, decreasing the time for credentialing is imperative. It is this problem that formed the impetus of the project.

SPCP group has many providers, in many different locations, with differing rates depending on the service(s) rendered. Keeping provider data current is challenging for RHI. SPCP requests monthly additions and updates to provider information. The SPCP group submits these changes to RHI in three ways: (1) carve-out list, (2) all employee list and (3) modifications based on a printed copy of the provider database. A carve out list is a list of medical services that are separated from a contract and paid under a different arrangement. Reports have overlaps in information and many people touch these reports before the department ultimately responsible for handling provider file maintenance receives them. For example, sometimes, the duplicate requests found on these three reports are inconsistent. Properly maintaining the provider file is important since the provider record drives the reimbursement the provider receives.

As a result of a poor process, the following impacts were felt by both RHI and the SPCP group: (1) system updates lagged well behind requests for updates due to the many "touches", (2) reimbursements to the providers were not timely nor accurate due to lag time mentioned in the first point, (3) inaccurate and untimely reimbursements result in almost 10,000 adjustments per year, (4) credentialing had to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Revenue management and provider enrollment cycles



perform duplicative tasks in order to complete their new provider process in a timely manner and (5) the SPCP group was very dissatisfied. This dissatisfaction led to the group withholding repayment of over $1 million in overpayments to induce RHI to reduce the time it took to make updates and to increase the accuracy of the provider file database itself.

The provider file data management (PFDM) department loads and maintains the provider information in all RHI systems. As mentioned before, the provider record drives provider reimbursement. PFDM processes service requests from SPCP, resulting in three main activities: the creation of new provider records, punching through the carve-out list, and making demographic updates. The analysis of the service request process requires that all three activities are examined.

The primary objective of this process improvement project was to reduce the time it takes to react to requests for changes from SPCP. Specifically, it addressed reducing the times associated with three activities: the creation of new provider records, punching through the carve-out list, and making demographic updates. The primary objective of the new provider creation process was to reduce the time it takes to create a new provider record from its current baseline of 55 calendar days to seven calendar days. The primary objective for punching through the carve-list was to reduce the time it takes from its current baseline of 40 calendar days to seven calendar days. Lastly, the primary objective for updating provider demographics was to reduce the time it takes from its current baseline of 50 calendar days to ten calendar days. Secondary, but important, benefits included reduced time to research the provider record requests and reduce the number of claim payment adjustment – resulting in an increase in SPCP's satisfaction.

By implementing a business intelligence system, the data was cleansed, multiple handoffs were eliminated, cycle time targets were met, and greater customer satisfaction was achieved. Specifically, the BI tool proactively identifies providers who were either (1) in the process of being credentialed, who had

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

not yet been approved, but that had started submitting claims, and (2) providers who were submitting claims but who had either not sent their credential application or that the application had been misplaced by RHI. In the first case, the BI tool was used to put a hold on all incoming claims, and sent an exception notice to the PFDM department indicating they needed to expedite the credentialing process. This substantially reduced the 10,000 adjustments per year. In addition, the BI tool was integrated with an existing credentialing database so better tracking of the approval process was available. The creation of the BI tool (and other improvements) allowed the company to exceed its stated objectives, improve SPCP's satisfaction, and decrease its costs. Next, we turn our attention to the Special Investigations Unit of RHI.

Case 2: Special Investigations Unit (SIU)

Increasing a Fraud Department's Ability to Detect Fraudulent Claims

Healthcare fraud and abuse affects everyone by increasing healthcare costs and healthcare coverage premiums. To combat fraud and abuse, the RHI's Special Investigations Unit (SIU) proactively seeks to detect and investigate potential healthcare fraud and abuse. SIU was experiencing an increase in the backlog of investigative leads. Under the leadership of its director, the department had a stated objective of "reducing the average investigation time (AIT) it takes for an investigator to complete an investigation by eliminating waste." To accomplish this objective, a two-pronged approach was pursued: (1) decrease the time for internal investigative procedures for assessing provider fraud, and (2) decrease the time for external investigations of member fraud. The first part involved a lean six sigma approach to streamline operations, and the second part involved active use of social media by investigators to seek out information on relationships and evidence that might indicate fraudulent behavior, coupled with the use of data analytics to connect fraudulent claims over time.

Fraud encountered by the company can be broken into two types – those frauds committed by providers and those acts committed by members. Common fraud schemes providers pursue include (1) billing for services that were never provided, (2) billing for services more expensive than those that were actually provided, (3) performing medically unnecessary services solely to receive insurance payments, (4) misrepresenting non-covered treatments as medically necessary covered treatments to receive payment, and (5) altering claims. Common frauds committed by members or individuals include (1) adding someone who is not eligible for insurance coverage or lending an ID card to family or friends, (2) submitting a fraudulent worker's compensation claim, and (3) submitting a fraudulent medical claim associated with staged accidents.

In Figure 5, the current state of the SIU's process is diagrammed. This process, though complete, would be more powerful with the addition of social media monitoring and a better BI tool to identify leads. The BI system would provide additional leads to be investigated, with supporting data already fed into the pre-audit review and internal audit procedures. Social media also supports CI, as multiple sources could report the same or similar relationships and red flags, allowing decision makers to come to quicker agreement. Overall, an effective system could be made more efficient with the addition of both BI and social media monitoring. Later in this section, mathematical equations will be used to show how BI and social media changes the balance of the equations so it is more costly for fraudsters to commit the crime.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Special investigations unit's current state

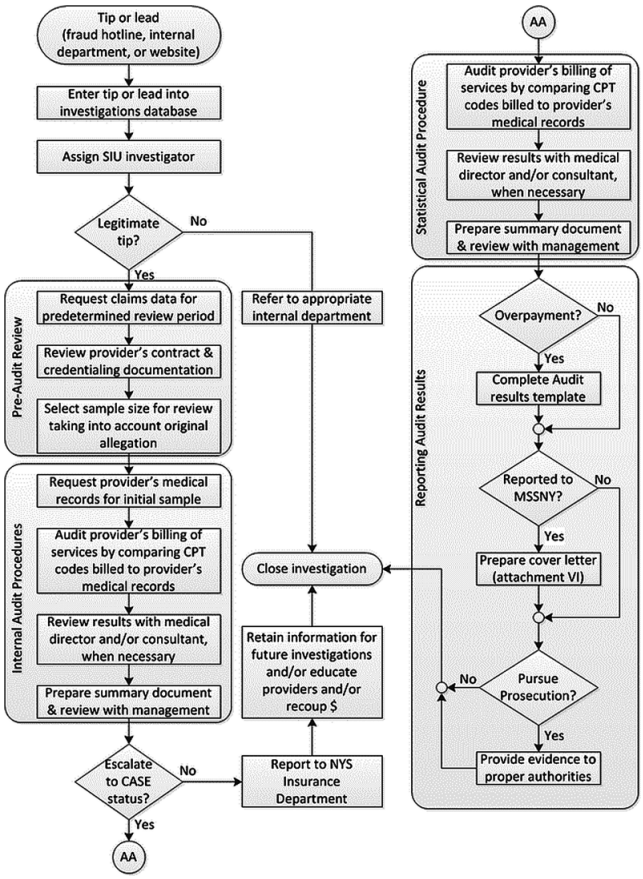
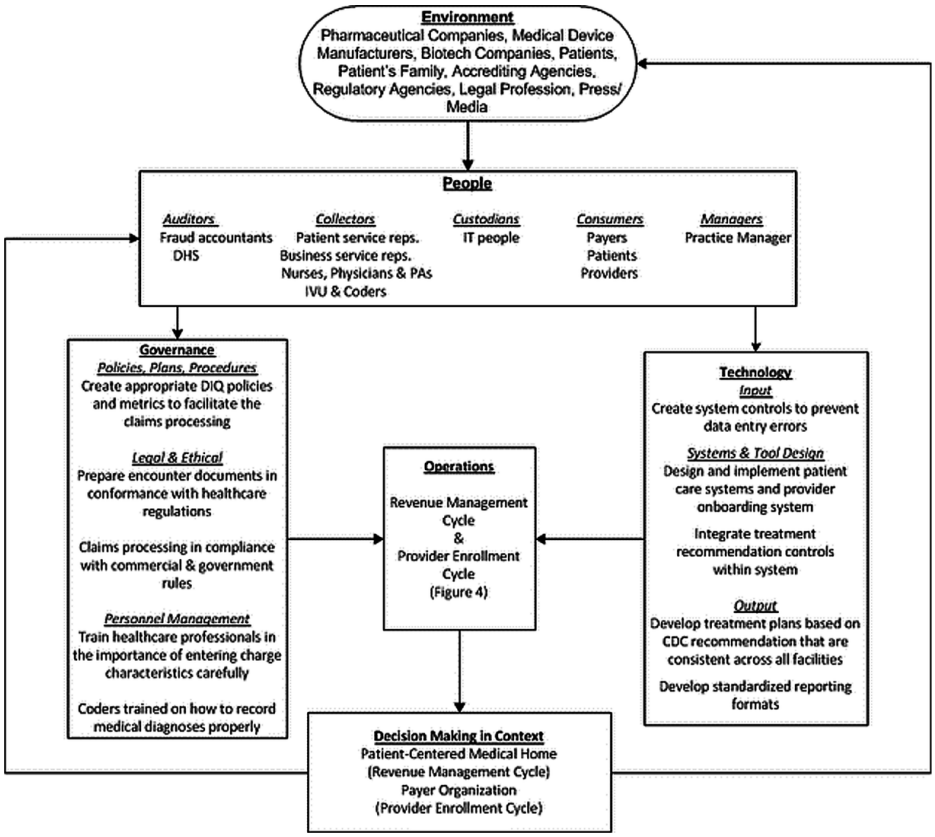


Figure 6 illustrates how the PGOT is used in a healthcare setting, with an emphasis on the people component of the model. A system implies components that work together, within a specific environment. In the healthcare realm the environment is broad and includes actors along the entire supply chain. These actors within the PGOT provide a framework to utilize collective intelligence. By definition, collective intelligence requires a consensus for any given action. By framing the actions within the governance, operations and technology components of the PGOT, the people can develop and implement policies and procedures for fraud investigation, operationalize the special investigations process and design the system for monitoring and reacting to the leads.

In stage 1 (decreasing the time for internal investigative procedures), when the team was gathered, it quickly surfaced that investigators were regularly and frequently answering calls from members concerning customer service issues that had nothing to do with fraud. In fact, what had happened was that customer service had eliminated their telephone number from the "Explanation of Benefits", leaving the fraud hotline number as the only number on the document. Members with legitimate customer service questions called the only number they had available – the fraud hotline. And that is exactly what happened – investigators were on the phones with members they should never have talked to in the first place. As a result of identifying this wasteful practice, the fraud hotline was relocated to the customer service department resulting in an improved member experience, increased satisfaction and quality, reduced wasted time, and decreased costs.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Roles with data governance



Next, we examine the financials of fraud from the fraudster's perspective, and the impact of actions by the payer on the costs. From this discussion, it will be clear that the development of BI and CI tools change the attractiveness of committing fraud, while only mildly increasing costs for payers.

Fraudster's and Company's Financial Perspective on Fraud

From the fraudster's perspective, the expected fraud cost (EFC) is a function of the cost of the criminal act (A), the cost of converting (if necessary) the misappropriated asset into something of value (C_1), and the cost of concealing the crime (C_2). To be clear, by definition, EFC consists of the costs associated with the fraudulent act itself, and not of the ramifications of being caught. The challenge in preventing healthcare fraud is that the expected cost of committing fraud for the fraudster is so small relative to the rewards one receives. For example, the conversion cost in healthcare fraud is practically zero since the fraudster receives money, and the sheer volume of claims in the system helps conceal the fraud itself. Therefore, very little effort is required by the fraudster, other than committing the act itself, to benefit from it. Equation (1) expresses these costs mathematically:

$$EFC\left(A, C_1, C_2\right) = A\left(P, D_1, D_2\right) + C_1 + C_2 \tag{1}$$

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The cost associated with being caught is explicitly addressed elsewhere in this analysis (see equation (2)). The cost of committing the fraudulent act (A), measured monetarily, is a function, among other things, of the prevention (P), deterrence (D_1), and detection (D_2) controls in place to stop such fraud from occurring. An example of A(P, D_1, D_2) is the cost associated with staging an accident in such a way that it is not obvious to investigators that it was staged. The more scrutiny individual accidents are given by investigators, the more effort criminals must put into staging it. The same holds true of providers who purposely code claims improperly in order to receive greater reimbursement dollars. Prevention, deterrence, and detection measures are designed to drive up the cost of committing the fraudulent act (e.g., A(P, D_1, D_2)) from the fraudster's perspective and the use of BI tools and social media monitoring are inexpensive ways to drive up a fraudster's cost. In general, the greater the controls, the greater the cost to the fraudster for planning and executing the commission of the crime.

In healthcare, prevention methods implemented by payers could include investigating whether a provider has any red flags concerning fraud before credentialing them. Another inexpensive preventative measure is sharing policies, plans, and protocols concerning fraud online with providers and members. Deterrence involves creating an ethical culture – consistent communication about the costs and punishments associated with fraudulent behavior, communicating a clear definition of what constitutes fraud, training healthcare professionals in the importance of accurately recording diagnoses, and training coders on how to record medical diagnoses properly. Detection mechanisms are also inexpensive ways to get good leads. These controls include tip hotlines, and fraud auditing.

The cost of converting what was misappropriated into something of value to the fraudster (measured monetarily) is almost always zero with health insurance fraud since the fraudster receives a check from the insurance company; therefore, the time depositing the check is all that is involved in the conversion process. In other industries, say where the misappropriated asset is inventory, the fraudster must spend time searching for a buyer for the inventory to convert the asset into cash. The cost of concealing the fraud, once again measured monetarily, depends on the circumstances surrounding the fraudulent act. At the very least, the fraudster has to consider the lost opportunity given their time could have been spent doing other, more productive things.

The expected fraud penalties (EFP) is a function of the probability (Pr) that the company's detection mechanisms (D_2) discovers the fraud, multiplied by the punishment cost (based on the company's response (R) and the value of the misappropriated asset (MA)), plus the negative emotional costs (E^-). The punishment costs (PC) include all formal and informal costs arising from punishment and lawsuits such as fines, lost income, and legal fees. Equation (2) expresses these costs mathematically:

$$EFP\left(E^{-}, D_2, R, MA\right) = E^{-} + \Pr\left(D_2\right) * PC\left(R, MA\right) \tag{2}$$

The negative emotional costs include guilt, anxiety, and fear to name a few; these costs are high(er) with crisis fraudsters, and low to nonexistent with career fraudsters. In terms of healthcare insurance fraud, most fraudsters are career fraudsters, and therefore are not burdened by the negative emotional costs that a crisis fraudster experiences.

The expected financial reward (EFR) received by the fraudster consists of two components: the positive emotional value (E+) from the criminal act itself, and the value from misappropriating the asset (MA), and if necessary, converting that asset into something else of value to the fraudster. The positive emotional value includes the thrill of danger, retribution, and sense of accomplishment to name a few;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

this value is high(er) for career fraudsters, and low to nonexistent with crisis fraudsters. No stand-alone equation is shown for EFR, but rather it is embedded within Equation (3).

The expected net profit (or loss) that a fraudster would receive if they committed the fraud repeatedly is denoted by ENP, and it equals the expected financial reward (EFR), minus both the expected fraud cost (EFC) and the expected fraud penalties (EFP). Equation (3) expresses these costs mathematically:

$$ENP = EFR\left(MA, E^{+}\right) - EFC\left(A, C_1, C_2\right) - EFP\left(E^{-}, D_2, R, MA\right) \tag{3}$$

Most of us would admit to having broken the law at some point in our lives; when was the last time you exceeded the speed limit or jaywalked? If you break the speed limit because you are late to work, then you are a "crisis" lawbreaker. On the other hand, if you regularly speed, run red lights, and do not yield to pedestrians, then you are a "career" lawbreaker. Each of us has a temptation threshold; this threshold is a function of many things, but an obvious parameter is our current income (prior to committing the fraud) and our moral compass. Therefore, even though all of us have different thresholds (and some individuals may claim their threshold is infinity), people follow this basic rule:

$$If\ \ ENP \geq temptation\ threshold,\ \ then\ commit\ the\ fraud;\ \ else,\ \ don't \tag{4}$$

The purpose of implementing BI tools and monitoring social media is to decrease the expected net profit to a value below the fraudster's temptation threshold. See Table 1 for a summary of the terms discussed above from the fraudster's perspective. Next, we will examine costs from the company's perspective, and then discuss how BI and social media can disrupt the cost equations.

The purpose of prevention, deterrence, detection, and response is to make sure that the cost equation is such that ENP is small enough that someone contemplating committing a fraud is not tempted to do so (assuming they are a rational decision maker and have a "normal" threshold). The use of business intelligence systems and social media increases the controls associated with prevention, deterrence, and detection at minimal costs, and hence, makes fraud less attractive.

If the company had no fraud strategy at all, then a certain percentage of the total assets (TA) would be misappropriated. The expected savings from having a fraud strategy (ESS) represents the difference between the estimated percentage of those assets that would have been stolen if no fraud strategy existed (L_{ns}) and the estimated percentage of those that actually were stolen as estimated by last period's losses given a strategy is in place and being executed (L_{ws}). Equation (5) expresses this relationship mathematically:

$$ESS\left(L_{ns}, L_{ws}, TA\right) = \left(L_{ns} - L_{ws}\right) * TA \tag{5}$$

The cost of a company's fraud strategy (CFS) is equal to the sum of the prevention, deterrence, detection, and response costs in a given time frame. The first three cost components of CFS were discussed earlier. Response (R) includes following the fraud response plan, gathering and presenting evidence, recommending punishments, and recovery of losses; these costs are denoted by C(R). Equation (6) expresses these costs mathematically:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Financials from the Fraudster's perspective

Acronym	Phrase	Description
A (P, D_1, D_2)	Cost of the Criminal Act	Cost of committing the fraudulent act is a function, among other things, of the prevention (P), deterrence (D_1), and detection (D_2) controls in place to stop such fraud
C_1	Conversion Costs	Cost of converting what was misappropriated into something of value to the fraudster (measured monetarily)
C_2	Concealment Costs	Cost of concealing the fraud (measured monetarily)
E^-	Negative Emotional Costs	These costs include guilt, anxiety, and fear to name a few
E^+	Positive Emotional Value	This value includes the thrill of danger, retribution, and sense of accomplishment to name a few
MA	Value of Misappropriated Asset	Value of misappropriated asset (e.g., time or more specifically labor costs, inventory, information) stolen by fraudster
PC	Punishment Costs	All formal and informal costs arising from punishment and lawsuits such as fines, lost income, and legal fees
EFR (MA, E^+)	Expected Financial Reward	Expected financial reward received by the fraudster from misappropriating the asset, and if necessary, converting that asset into something else of value to the fraudster, plus any positive emotional value experienced
EFC (A, C_1, C_2)	Expected Fraud Cost	Function of the cost of the criminal act (A), the cost of converting the misappropriated asset into something of value (C_1), and the cost of concealing the crime (C_2)
EFP (E^-, D_2, R, MA)	Expected Fraud Penalties	A function of the probability (Pr) that the company's detection mechanisms (D_2) discovers the fraud multiplied by the punishment cost (based on the company's response (R) and the value of the misappropriated asset (MA)) plus the negative emotional costs (E^-)
ENP	Expected Net Profit	The expected net profit (or loss) that a fraudster would receive if they committed the fraud repeatedly; it equals the expected financial reward (EFR) minus both the expected fraud cost (EFC) and the expected fraud penalties (EFP)

$$CFS\left(P, D_1, D_2, R\right) = C\left(P\right) + C\left(D_1\right) + C\left(D_2\right) + C\left(R\right) \tag{6}$$

The expected net financial impact (NFI) on the company of having a fraud strategy in place and implementing it is equal to the expected savings from having a fraud strategy (ESS), minus the cost of implementing the company's fraud strategy (CFS). Equation (7) expresses this relationship mathematically:

$$NFI = ESS\left(L_{ns}, L_{ws}, TA\right) - CFS\left(P, D_1, D_2, R\right) \tag{7}$$

Without a fraud strategy, fraud skyrockets. Therefore, ENP and NFI are related. The fraud strategy pursued by a company impacts both the EFC and EFP, which determines the ENP. And the average size of ENP impacts the need for a fraud strategy (and its associated costs). Change the fraud strategy (and NFI), and you change the behavior of the fraudster (and their ENP):

$$If\ NFI > risk\ threshold\left(\$\right),\ then\ execute\ a\ fraud\ strategy;\ else,\ don't \tag{8}$$

We have presented these equations so the reader can understand how the costs incurred by the fraudster are impacted by or a function of the controls put in place by the company (see Table 2).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Financials from company's perspective

Acronym	Phrase	Description
P	Prevention	Preventative controls include creating an anti-fraud policy, fraud audit plan and response protocol, and a fraud response plan; the costs of putting in place preventative controls is denoted by C(P)
D_1	Deterrence	Deterrence involves creating an ethical culture; the costs of putting in place deterrence mechanisms is denoted by $C(D_1)$
D_2	Detection	Detection mechanisms include tip hotlines, and fraud auditing; the costs of mechanisms designed to uncover fraud is denoted by $C(D_2)$
R	Response	Response includes following the fraud response plan, gathering and presenting evidence, recommending punishments, and recovery of losses; these costs is denoted by C(R)
CFS (P, D_1, D_2, R)	Cost of Fraud Strategy	The cost of a company's fraud strategy is equal to the sum of the prevention, deterrence, detection, and response costs in a given time frame.
L_{ns}	Loss Factor (no strategy)	Estimated percentage of assets that would be misappropriated without a fraud strategy in place
L_{ws}	Loss Factor (with strategy)	Estimated percentage of assets that would be misappropriated even though the company has a fraud strategy in place and implements it
TA	Total Assets	Sum of the value of all assets that the company cares to protect from fraud and respond to if misappropriated
ESS (L_{ns}, L_{ws}, TA)	Expected Savings from (Fraud) Strategy	If the company had no fraud strategy, then a certain percentage of assets would be misappropriated; the expected savings from having a fraud strategy represents the difference between those assets that would have been stolen if no fraud strategy existed and those that actually were (as estimated by last period's losses)
NFI	Net Financial Impact	Expected net financial impact on the company of having a fraud strategy in place and implementing it; it equals the expected savings from fraud strategy, minus the cost of the company's fraud strategy

Figure 7 shows the relationship between the work environment and the social environment. By using social media, the company is able to quickly and cost effectively look for red flags and relationships among claimants or providers that might indicate fraud exists. Next, we examine some very specific recommendations for action to improve process performance with minimal costs.

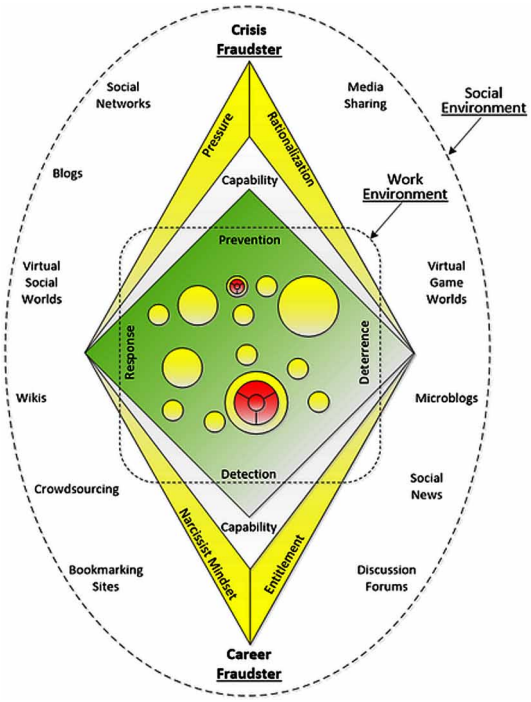
RECOMMENDATIONS AND MODEL VALIDATION

BI systems have the potential to decrease costs in the healthcare industry by improving DIQ. Such systems provide information that will enable better patient care and help healthcare organizations survive with expedited (but smaller) payments, and decreasing fraud. However, these benefits will only accrue if the data that drives the information is of high quality and is appropriate for the decisions being made.

The PGOT framework, with its focus on the systemic effects of DIQ, can provide structure to the development and use of a BI system in healthcare. Thoughtful consideration of the implications of policies and procedures with respect to DIQ, appointing appropriate people in key roles such as Chief Data Steward and a Governance Council, and ensuring that systems are operationally and technically aligned with the DIQ goals, can help ensure that the BI systems can provide the information needed by payers to efficiently credential and process claims, and investigate frauds.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Fraud compass with social media impact



As seen in both cases, one of the difficulties facing healthcare is the potential time involved in performing processes, whether credentialing or investigating fraud. Given the size and scope of current healthcare organizations, it is common to find long, drawn out processes. Figure 6 illustrates how the PGOT can be applied to healthcare, specifically with respect to fraud investigation, a key driver for ensuring the financial success of the health insurance organization. By examining the BI system with respect to DIQ systemically, we can help ensure that the organization as a whole is running at optimal efficiency. For example, DIQ governance in healthcare can facilitate claims processing, credentialing, and special investigations. As indicated earlier, the healthcare industry does not deal with one payer and one payee, but rather multiple payers and payees. As seen in Figure 4, the revenue cycle management can only be enabled if the providers are successfully enrolled in the system with the right credentials and privileges. Without providers, there would be nothing to bill for. Without billing, the healthcare organization would not be able to support itself.

Thus, our recommendation is that a multi-step approach be followed:

1. Create a high-level context diagram that shows the system connections to the environment and closely examine the roles of people as they apply to the process (see Figures 4 and 6 for examples);
2. Create a process flow chart that shows the stakeholders, activities and system data flows in question (see Figure 5);
3. Align the PGOT framework (Figure 3) to the process flow chart to identify areas of concern (Figure 5);
4. Implement interventions to close the gap between the "as is" and the "to be";

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. For special investigations, use social media to investigate relationships and red flags (see Figure 7);
6. Track DIQ and process performance, continuing to make adjustments as needed.

IS design research emphasizes new method development over the use and evaluation of existing methods (Moody, 2003). There is a real lack of empirical research validating methods in practice, and a need to apply methods by practitioners within real organizational contexts. There are inherent problems evaluating any methodology or design technique given typically no theory, hypothesis, experimental design or data analysis exists to apply traditional evaluation criteria (Weber 1997). There are many more IS design methods published in the research literature than what has been implemented or even tested in practice. (Fitzgerald, 1991) persuasively makes the case that usage is an important pragmatic measure of the "success" of a method, and should be considered when evaluating research.

Moody (2003) combined the Technology Acceptance Model (TAM) from the IS literature, and Methodological Pragmatism from the philosophy of science to construct an approach for evaluating IS design methods. There are two dimensions of success: Actual efficacy which is defined as whether the method improves tasks performance, and adoption in practice of the method regardless of whether the method improves performance or not. This method evaluation model developed by Moody (2003) validates our model since our approach both improved task performance (i.e., actual efficacy) and was adopted in practice. With this methodical approach to examining the problem and solution, with a careful examination of the DIQ issues along the way, the BI system can be implemented to take full advantage of the power that it holds.

CONTRIBUTIONS TO THE LITERATURE

This paper contributes to the literature in healthcare, fraud management, and business intelligence. A new model of fraud is proposed: the fraud compass. Using this model and the PGOT model introduced in Neely and Cook (2011), we demonstrate how social media can be used in the early detection of fraud. Additionally, the PGOT model is implemented in a specific environment, showing that the model can move from the general to the specific. The development of mathematical models to support fraud management are discussed. Further research is needed to test the applicability of these mathematical models of fraud in other settings. Finally, practitioner recommendations are presented, with an overarching goal of improving healthcare processes, through diagrams and implementation. Future research can focus on implementing the PGOT in other domains, exploring the relevance of the fraud compass and developing experiments to test the validity of the mathematical models.

CONCLUSION

This paper describes two case studies in the healthcare industry. By weaving the elements of the PGOT framework throughout the narrative, several opportunities for improvement were suggested and discussed. This led to recommendations for best practices. This paper adds to the theoretical literature by operationalizing a theoretical model and showing the interconnections within an overall DIQ system. It adds to the practical literature by offering specific recommendations for practice. DIQ is obviously

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

an important issue in healthcare. It affects the ability of providers and business managers at all levels to make decisions. Within the healthcare industry there are many opportunities for improved decision making, which can directly impact medical outcomes and the bottom line.

There are many organizational systems including transaction processing systems (TPS), management information systems (MIS), enterprise resource planning systems (ERP) and business intelligence (BI) systems. As discussed in the literature review, BI is not just a technology. It is the system which collects, stores, and analyzes business data to support decision making. Key components of this system include source systems, the data warehouse and associated metadata, and the analytical tools such as data mining and general statistical analysis. Moving data from the sources to the data warehouse requires data integration which is both a technology and a process. Metadata, data quality and data governance processes ensure that data moves through the system efficiently and effectively (Wixom & Watson, 2010). Integrating DIQ methodologies into the BI system will help to ensure that the system addresses the end user needs. The descriptions of the cases in this article show how this can be accomplished. CI can support fraud investigation, through the use of social media monitoring. Taken together, DIQ, BI and CI will improve all aspects of the healthcare system, with the two case studies showing examples of where these techniques can be used.

REFERENCES

Agency for Healthcare Research and Quality. (2012). Patient Centered Medical Home. Retrieved from http://pcmh.ahrq.gov/portal/server.pt/community/pcmh__home/1483/PCMH_Defining%20the%20PCMH_v2

Bardhan, I. R., & Thouin, M. F. (2012). Health information technology and its impact on the quality and cost of healthcare delivery. *Decision Support Systems*, *55*(2), 438–449. doi:10.1016/j.dss.2012.10.003

Bitterer, A., Schlegel, K., Hostmann, B., & Gassman, B. Beyer, MA, Herschel, G, … Andrews, W. (2007). Hype Cycle for Business Intelligence and Performance Management. Stamford, CT: Gartner Research.

Bonchi, F., Giannotti, F., Mainetto, G., & Pedreschi, D. (1999, August). A classification-based methodology for planning audit strategies in fraud detection. *Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 175-184). ACM. 10.1145/312129.312224

Boritz, J. E. (2005). IS practitioners views on core concepts of information integrity. *International Journal of Accounting Information Systems*, *6*(4), 20. doi:10.1016/j.accinf.2005.07.001

Bosler, B. (2015). An interview with deputy chief, health care fraud unit at the US Attorney's office. *Home healthcare now, 33*(10), 560-562.

Bourquard, J., & Kirsch, C. (2014). Big data= big benefits: States are using large amounts of data to improve efficiency, fight fraud, and identify savings. *State Legislatures*, *40*(8), 32–34. PMID:25241420

Byrd, L. W., & Byrd, T. A. (2012). The impact of clinical information technology on information quality in US hospitals. *International Journal of Healthcare Technology and Management*, *13*(1), 71–87. doi:10.1504/IJHTM.2012.048948

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chang, V. (2014). The business intelligence as a service in the cloud. *Future Generation Computer Systems*, *37*, 512–534. doi:10.1016/j.future.2013.12.028

Chang, V. (2015). A cybernetics social cloud. *Journal of Systems and Software.*

Chen, H., Chiang, R.H.L., & Storey, V.C. (2012). Business intelligence and analytics: From big data to big impact. *Management Information Systems Quarterly*, *36*(4), 1165–1188.

Davenport, T.H. (2010). Business intelligence and organizational decisions. *International Journal of Business Intelligence Research*, *1*(1), 1–12. doi:10.4018/jbir.2010071701

Debreceny, R.S., Gray, G.L., Ng, J.J.J., Lee, K.S.P., & Yau, W.F. (2005). Embedded audit modules in enterprise resource planning systems: Implementation and functionality. *Journal of Information Systems*, *19*(2), 7–27. doi:10.2308/jis.2005.19.2.7

DeLancey, M., & Gorton, L. (2015). Medicare Advantage Fraud: Temptation, Consequence, and Protection. *Managed Care (Langhorne, Pa.)*, *24*(11), 57. PMID:26672172

Dorminey, Jack, Fleming, A.S., Kranacher, M.-J., & Riley, R.A., Jr. (2012). The evolution of fraud theory. *Issues in Accounting Education,* 27(2), 555-579.

Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the mind in the eyes or reading between the lines? Theory of mind predicts collective intelligence equally well online and face-to-face. *PLoS ONE*, *9*(12), e115212. doi:10.1371/journal.pone.0115212 PMID:25514387

Fawcett, T., & Provost, F. (1997). Adaptive fraud detection. *Data Mining and Knowledge Discovery*, *1*(3), 291–316. doi:10.1023/A:1009700419189

Fitzgerald, G. (1991). *Validating New Information Systems Techniques: A Retrospective Analysis. In* H.E. Nissen, H.K. Klein, & R. Hirschheim, (Eds.), *Information Systems Research: Contemporary Approaches and Emergent Traditions*.

Ge, M., & Helfert, M. (2008). Effects of information quality on inventory management. *International Journal of Information Quality*, *2*(2), 15. doi:10.1504/IJIQ.2008.022962

Giannangelo, K., & Fenton, S. (2008). EHR's effect on the revenue cycle management coding function. *Journal of Healthcare Information Management*, *22*(1), 26–29. PMID:19267004

Harmon, K. (2009). *Deaths from avoidable medical error more than double in past decade, investigation shows*. Scientific American Blog.

Hernández-Chan, G., Rodríguez-González, A., Alor-Hernández, G., Gómez-Berbís, J. M., Mayer-Pujadas, M. A., & Posada-Gómez, R. (2012). Knowledge acquisition for medical diagnosis using collective intelligence. *Journal of Medical Systems*, *36*(1), 5–9. doi:10.100710916-012-9886-3 PMID:23089960

Imhoff, C. (2005). Streaming Processes for Front-Line Workers: Adding Business Intelligence for Operations. Operational Business Intelligence White Paper. Retrieved from http://www.tdwi.com

Indart, B. (2006). Operationalizing business intelligence – turning insight into actions. *Business Intelligence Journal*, *11*(2), 35.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jane Budge, E., Maria Tsoti, S., James Howgate, D., Sivakumar, S., & Jalali, M. (2015). Collective intelligence for translational medicine: Crowdsourcing insights and innovation from an interdisciplinary biomedical research community. *Annals of Medicine*, *47*(7), 570–575. doi:10.3109/07853890.2015.1091945 PMID:26469375

Kerr, K. A., Norris, T., & Stockdale, R. (2008). The strategic management of data quality in healthcare. *Health Informatics Journal*, *14*(4), 259–266. doi:10.1177/1460458208096555 PMID:19008276

Kurvers, R. H., Krause, J., Argenziano, G., Zalaudek, I., & Wolf, M. (2015). Detection Accuracy of Collective Intelligence Assessments for Skin Cancer Diagnosis. *JAMA Dermatology*. PMID:26501400

Linder, J. A., Ma, J., Bates, D. W., Middleton, B., & Stafford, R. S. (2007). Electronic health record use and the quality of ambulatory care in the United States. *Archives of Internal Medicine*, *167*(13), 1400–1405. doi:10.1001/archinte.167.13.1400 PMID:17620534

Maguire, H. (2008). Extending the boundaries of IQ: Can collaboration with information management improve corporate governance? *International Journal of Information Quality*, *2*(1), 23. doi:10.1504/IJIQ.2008.019561

Marjanovic, O. (2010). The Importance of Process Thinking in Business Intelligence. *International Journal of Business Intelligence Research*, *1*(4), 29–46. doi:10.4018/jbir.2010100102

Mj, Romano, Stafford R. S. (2011). Electronic health records and clinical decision support systems: Impact on national ambulatory care quality. *Archives of Internal Medicine*, *171*(10), 897–903. doi:10.1001/archinternmed.2010.527 PMID:21263077

Moody, D.L. (2003). The Method Evaluation Model: A Theoretical Model for Validating Information Systems Design Methods. *Proceedings of ECIS '03.* Retrieved from http://aisel.aisnet.org/ecis2003/79

Neely, M. P., & Cook, J. S. (2011). Fifteen years of data and information quality literature: Developing a research agenda for accounting. *Journal of Information Systems*, *25*(1), 79–108. doi:10.2308/jis.2011.25.1.79

Olinsky, A., & Schumacher, P. A. (2010). Data Mining for Health Care Professionals: MBA Course Projects Resulting in Hospital Improvements. *International Journal of Business Intelligence Research*, *1*(2), 30–41. doi:10.4018/jbir.2010040104

Pighin, M., & Ieronutti, L. (2008). A Methodology Supporting the Design and Evaluating the Final Quality of Data Warehouses. *International Journal of Data Warehousing and Mining*, *4*(3), 20. doi:10.4018/jdwm.2008070102

Rauscher, S., & Wheeler, JR. (2008). Effective hospital revenue cycle management: Is there a tradeoff between the amount of patient revenue and the speed of revenue collection? *Journal of Healthcare Management/American College of Healthcare Executives, 53*(6), 392.

Rauscher, S., & Wheeler, J. R. C. (2010). Hospital Revenue Cycle Management and Payer Mix: Do Medicare and Medicaid Undermine Hospitals' Ability to Generate and Collect Patient Care Revenue? *Editorial Board*, *37*(2), 90–104. PMID:21294440

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sillup, G. P., Klimberg, R. K., & McSweeney, D. P. (2010). Data-Driven Decision Making for New Drugs: A Collaborative Learning Experience. *International Journal of Business Intelligence Research*, *1*(2), 42–59. doi:10.4018/jbir.2010040105

Simborg, D. W. (2014). Fraud And Electronic Health Records. *Health Affairs*, *33*(10), 1877–1877. doi:10.1377/hlthaff.2014.0916 PMID:25288434

Sunyaev, A., & Chornyi, D. (2012). Supporting chronic disease care quality: Design and implementation of a health service and its integration with electronic health records. *Journal of Data and Information Quality*, *3*(2), 1–21. doi:10.1145/2184442.2184443

U.S. Department of Health and Human Services. (2001). Final Rules for Administrative Simplification. Retrieved from http://aspe.hhs.gov/admnsimp/final/PvcPre02.htm

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All---A Contingency Approach to Data Governance. *Journal of Data and Information Quality*, *1*(1), 4. doi:10.1145/1515693.1515696

Wickramasinghe, N., & Schaffer, J. L. (2006). Creating knowledge-driven healthcare processes with the intelligence continuum. *International Journal of Electronic Healthcare*, *2*(2), 164–174. doi:10.1504/IJEH.2006.008830 PMID:18048242

Wixom, B. H., & Watson, H. J. (2010). The BI-Based Organisation. *International Journal of Business Intelligence Research*, *1*(1), 13–28. doi:10.4018/jbir.2010071702

Wolf, M., Krause, J., Carney, P. A., Bogart, A., & Kurvers, R. H. (2015). Collective Intelligence Meets Medical Decision-Making: The Collective Outperforms the Best Radiologist. *PLoS ONE*, *10*(8), e0134269. doi:10.1371/journal.pone.0134269 PMID:26267331

This research was previously published in the International Journal of Organizational and Collective Intelligence (IJOCI), 7(2); edited by Victor Chang; pages 13-33; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 89
Forensic Investigation of Digital Crimes in Healthcare Applications

Nourhene Ellouze
University of Carthage, Tunisia

Slim Rekhis
University of Carthage, Tunisia

Noureddine Boudriga
University of Carthage, Tunisia

ABSTRACT

Healthcare applications are increasingly being used due to the safety and convenience brought to patients' life and healthcare professionals, respectively. Nevertheless, the use of weak authentication techniques and vulnerable communication protocols makes these applications threatened by specific classes of security attacks and e-crimes. The latter threaten the privacy, the safety and even the life of the persons using these applications, due to the fact that they handle sensitive information and implement complex and critical features. This chapter focuses on postmortem investigation of crimes on healthcare applications. After classifying crimes targeting healthcare applications, the requirements for the design of appropriate postmortem investigation system, are discussed. A literature review of proposals related to the investigation of crimes in healthcare applications together with a discussion of the advanced issues are also provided in this chapter.

INTRODUCTION

The advances in Information and Communications Technologies have led to the release of a wide set of healthcare applications, including, but are not limited to, remote monitoring of patients in hospitals, real time detection of emergency situations threatening chronically ill patients, and continuous monitoring of

DOI: 10.4018/978-1-7998-1204-3.ch089

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

elderly people. Healthcare applications are increasingly being used owing to the great efficiency and quality of service they offer to elderly persons, patients, and healthcare professionals. They have contributed to the enhancement of patients' autonomy, and the improvement of clinical treatment and remote health surveillance through the development and implementation of several technologies and equipment like, Wireless Sensor Networks (WSN), Implantable Medical Devices (IMDs), and cloud services.

Healthcare applications exhibit significant security weaknesses including the use of weak authentication techniques and the lack of appropriate security mechanisms. These weaknesses make them unprotected and subject to several criminal attacks threatening the safety and the privacy of patients, especially as these applications handle sensitive medical data and provide life-sustaining functions. Among the most common crimes targeting healthcare applications, we cite the unauthorized access to the medical records. Such type of attacks may induce serious threats on the privacy of the victims through the disclosure of their medical records, or even may threaten their life through the malicious modification of their medical records to make further medical prescription by physicians erroneous. In this context, the design of a system for postmortem investigation of criminal incidents threatening healthcare systems is becoming a key requirement.

A set of challenging issues should be addressed during the design of a postmortem investigation system tailored to healthcare applications. The first challenge is related to the storage of the huge amount of evidential traces that can be provided by healthcare systems, as these traces require an unlimited storage space when collected over a long period of time. The second challenge is related to the integrity and the trustworthiness of the provided evidence, especially as healthcare systems are connected to open environments (e.g., internet) through vulnerable communications protocols. The third is related to the complexity of medical evidence collection and processing, especially as healthcare systems implement different technologies and equipment that may provide heterogeneous evidence. The fourth challenge is related to the need that an investigation on healthcare digital crimes collects and analyzes two types evidence collected from the IT system under investigation. The first type contains information related to authentication, access, sensitive events execution, while the second type contains the medical information related to the victim health status.

Several research works have focused on the identification of criminal attacks targeting healthcare applications, and the development of security techniques protecting them from security attacks. Other research works focused on the development of secure audit and accounting tailored to healthcare applications to prepare for a postmortem investigation. A set of postmortem investigation techniques were also proposed in the literature to address medical and physical crimes. An overview of these research works will be provided in this chapter.

The remaining part of this chapter is organized as follows. The next section provides an overview of e-health systems. In Section 2, we identify and classify crimes targeting healthcare systems. Section 3 discusses the main requirements for the design of postmortem investigation of e-crimes on healthcare applications. In Section 4, we review the literature related to the design of techniques and methodologies for the checking of audit logs and the analysis of incidents on Wireless Body Area networks. We also discuss in this section digital investigation scope and techniques considering attacks on Implantable Medical Devices. Section 5 reviews investigation techniques tailored to physical crimes. In Section 6, we discuss advanced issued related to the investigation of criminal attacks on healthcare applications.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

SECTION 1: OVERVIEW OF E-HEALTH SYSTEMS

E-health systems refer to the systems providing healthcare services (e.g., diagnosis, treatment, monitoring) using Information and Communication Technologies (ICTs). These systems can largely serve patients safety by making healthcare services more efficient and improving the work of healthcare professionals (e.g., physicians and caregivers). There are several types of healthcare applications such as remote monitoring of patients, and real time detection of emergency situations which can improve the quality of life of elderly and chronically ill persons, as they reduce the amount of travelling (e.g., to book diagnosis appointments) and avoid hospitalization. In the remaining part of this section, an overview of healthcare applications will be provided.

Generic Architecture of E-Health Systems

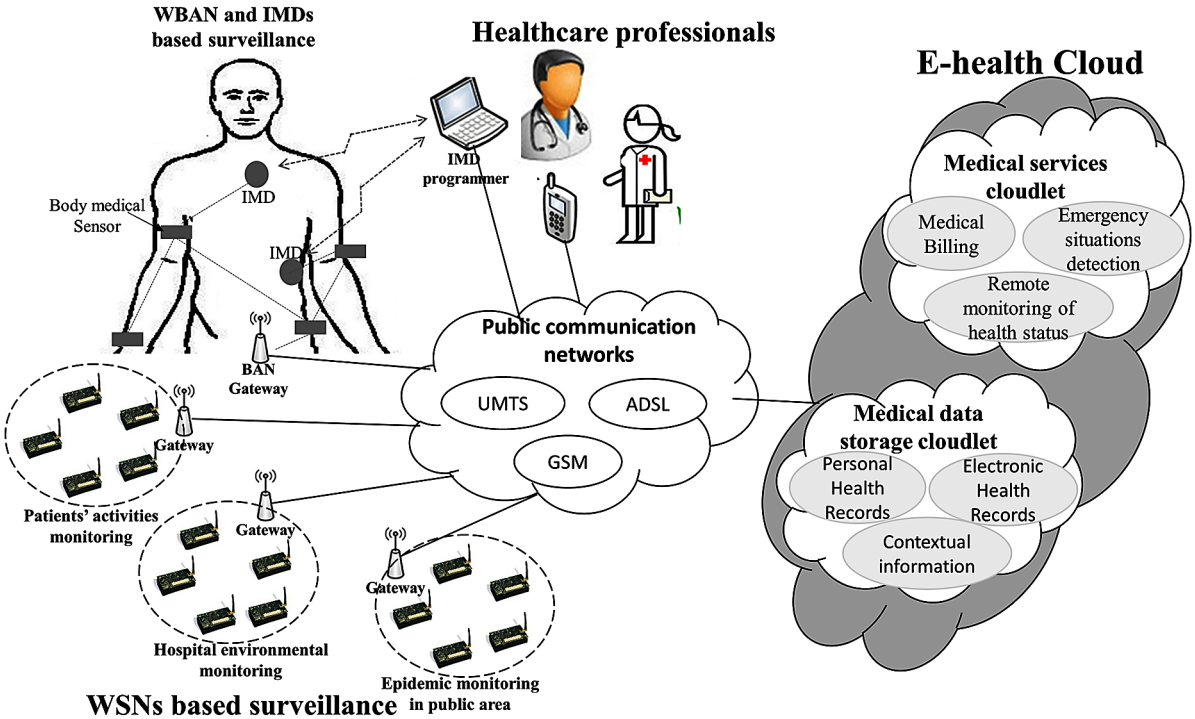
Typically, e-health systems involve three principal stakeholders: a) the healthcare service providers, which are responsible of managing and providing healthcare services; b) the patient, who represents the person eligible for the provided healthcare services; and c) the healthcare professionals (e.g., physicians and caregivers) which use e-health systems to perform their healthcare functions, such as the remote diagnosis of patients.

Several architectures of e-health systems were proposed in the literature. They vary according to the type of the designed healthcare applications, the provided services, and the implemented technologies. In the sequel, we provide a generic architecture of e-health systems that is able to implement the majority of applications and technologies. Figure 1 illustrates the generic architecture of an e-health system which mainly integrates the following described components.

- **Wireless Sensor Network (WSN):** Consists of a set of environmental sensor nodes including, but are not limited to video cameras, gas, pressure, temperature, humidity, and locations sensors. These nodes are typically deployed around the monitored person (e.g., home, hospital) to provide rich contextual information about him and identify the environmental conditions where he is located. Nodes of the same network communicate through a multi-hop wireless technology (e.g., ZigBee) to forward the sensed data to the sink node or the gateway.
- **Wireless Body Area Network (WBAN):** Represents a specific type of Wireless Sensor Network, which consists of a set of medical sensor nodes (e.g., physiological, cardiac, and glucose sensors) that are deployed on the body of a person to monitor its vital signs. These nodes can be interconnected through a wireless network by means of different technologies, such as IEEE 802.15.4. Body medical nodes belonging to the same WBAN cooperate together to forward the performed measurements to the BAN gateway, which will be in charge of transmitting these data through a public communication network to be stored in the corresponding medical data storage cloud/cloudlet.
- **Implantable Medical Device (IMD):** Is a medical device that is surgically implanted inside the body of a patient who suffers from chronic ailments, such as cardiac arrhythmia, diabetes, and Parkinson's disease. Such a device is in charge of measuring the vital signs of the patient (e.g., heart rate, glucose level, or blood pressure) and applying specific medical therapeutic functions when abnormal physiological conditions appear. Because they are surgically implanted, IMDs require a wireless connection to allow healthcare professionals to collect of data stored in them (e.g.,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Generic architecture of e-health systems



the history of the patient's vital signs and the history of responses undertaken by the IMD), monitor in real time the patient's vital signs, adjust the therapy configuration, and update the device firmware. To this end, IMDs can be wirelessly configured using an external programmer through the 402-405 MHz Medical Implant Communication Service (Savci et al., 2005) band. Thanks to the integration of IEEE 802.15.4 and 6LoWPAN (Hui & Culler, 2008) technologies, future IMDs would also be wirelessly connected to the public communication network (Astorga et al., 2014) through a BAN gateway.

- **Gateway:** Ensures the adaptation layer between the Internet Protocol (IP) stack's link, and Wireless Body Area and Sensor networks, so that IP packets could be sent and received over these networks. A gateway is in charge of forwarding the data received from WBANs or WSNs through the public communication network to the medical data storage cloudlet to be stored. It is also responsible of transmitting the data requests received from healthcare professionals to the corresponding entity (i.e., WSNs or WBAN).
- **Public Communication Networks:** They ensure the interconnection between the WSNs (as well as the WBAN) part of the e-health system, the healthcare professionals, and different used cloudlets. Several IP and packet switched and cellular networks can be used to offer a fixed or mobile broadband access. The selection of the communication technologies to be used depends on the specific requirements regarding the quality of service and security to be provided. In healthcare applications, where medical equipment need to transfer patients' medical using wireless communication, while reducing communication interference, prioritizing access to the wireless channel by remote medical equipment, protect medical services from unavailability and harmful interferences, cognitive radio networks can be provided as a solution for e-health applications in hospitals.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **E-Health Cloud:** There are several types of clouds which are designed to provide a large scale of services and applications. Among the most common types of cloud services, we cite: Cloud Storage as a Service (StaaS) and Cloud Software as a Service (SaaS). StaaS ensures an infinite remote storage of data and allows users to access them any time from any place. It also allows cloud applications to share data between them and scale beyond their limited servers. SaaS makes available a set of applications running on a cloud infrastructure and accessible to the user through a thin client interface without requiring to manage or control the underlying cloud infrastructure (e.g., network, storage, and operating systems).

As depicted in Figure 1, two main types of interaction scenarios may be distinguished: regular and emergency. While communication in the first scenario is periodically performed by the e-health system, the second scenario is only used in emergency situations (when the healthcare professionals notice a deterioration of the health status of the patient, and need to get a real-time data information to identify the actual health status of the patient).

In the regular scenario, the WBAN and the deployed WSNs collect periodically a set of medical and contextual information. After that, the collected information is forwarded by the gateways of these networks to the medical data storage cloud/cloudlets through public communication networks. After being stored in the medical data storage cloud/cloudlet, these data will be accessible to the healthcare professionals. Generally, healthcare professionals use third-party applications to get access to/analyze the data related to the supervised patients.

In the emergency scenario, which is triggered by the detection of bad health parameters, the healthcare professionals need to collect instantaneous information about the patients' health status, and be able to remotely order immediate medication, intervention, or emergency ambulance transportation. All the emergency requests exchanged with the different healthcare professionals and the WSNs/WBANs should be timely transmitted via the communication system, while ensuring a high level of quality of service to control the end-to-end delay and network throughput. Therefore, at least two levels of quality need to be provisioned by the network depending in the interaction scenario: regular, or emergency.

Wireless Sensor Networks for Healthcare

Thanks to the continuous monitoring of the vital functions that can be offered by sensor nodes, WSNs for healthcare applications have the potential to offer a set of benefits to patients, healthcare professionals, and the society. First, they provide a continuous long-term monitoring of chronically ill patients, confirm the adherence of these patients to treatment guidelines, and monitor the effects of drug therapy. Second, they provide a real time feedback of the health status of patients. Third, they allow the early detection of life-threatening and emergency situations, so that a proactive reaction could be undertaken by the medical staff. Fourth, WSNs for healthcare applications allow the collection of context aware information (e.g., location, motions, and activities) which are useful to identify unusual patterns of the monitored persons. Fifth, they provide the early detection of serious external threats on patients' safety such as air pollution and presence of virus, as they are able to perform a continuous monitoring of a wide set of environmental and physiological parameters. Because of the benefits offered by WSNs for healthcare, a set of WSN based healthcare applications was proposed in the literature. Among these applications, we cite:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. Tracking and monitoring of patients inside hospitals;
2. Monitoring the physiological parameters of chronically ill persons; and
3. Surveillance of the behavior and activities of children and elderly peoples.

A set of prototypes of WSN based healthcare systems was also designed to monitor the health status of patients and transmit his/her physiological data. Among them, we cite:

1. Medical Emergency Detection in Sensor Network (Ko et al., 2010) system. Such a system ensures the monitoring of the physiological parameters of patients inside hospitals;
2. Medical Ad hoc Sensor Network (Hu et al., 2008) system which is in charge of ensuring the real-time surveillance of patients who suffer from heart diseases; and
3. A body area network based system (Felisberto et al., 2015) for the recognition of human activities and the detection of harmful events.

Several issues should be addressed during the design of WSNs for healthcare applications. One of the principal issues is the preservation of power consumption, since sensor nodes use a battery with limited lifetime whose replacement is a burdensome task. To this end, it is required to extend the availability of sensor nodes by minimizing to the maximum possible the energy overhead induced by the implemented algorithms. Privacy and security are also among the critical issues that should be addressed. In fact, the unauthorized collection and use of the sensitive data collected by sensor nodes can cause life threatening risks to the proprietary patient.

Implantable Medical Devices for Healthcare

Implantable Medical Devices (IMDs) are increasingly being used by chronically ill patients, due to the benefits, the life-sustaining, and the therapeutic functions they offered. In fact, they have improved the quality of life of patients by making it possible the continuous control of their health status and the automatic delivery of therapy when abnormal physiological conditions appear. Pacemakers, for example, are implanted in the heart tissue of a patient to regulate its heart rhythm through the delivery of electric shocks.

A generic architecture of an IMD mainly integrates:

1. A set of sensing devices to measure the physiological parameters useful to identify the health status of the patient;
2. A processing unit to manage the functionalities of the device;
3. A stimulator to deliver the required therapy;
4. A memory to store the collected measurements and therapy settings;
5. A transceiver to enable the wireless communication with the programmer used by physicians; and
6. A battery which is in charge of powering the device components.

The principal issues of IMDs are the preservation of energy consumption and the guarantee of the privacy and the security of the patients carrying them. In fact, an IMD should remain operational as long as possible, despite the fact that it uses a limited battery whose replacement requires a surgical intervention. In addition, due to the use of wireless communication, IMDs are threatened by a set of se-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

curity attacks, where most of them could be a serious source of death of patients carrying these devices. For instance, since an IMD automatically reacts to treat the detected cardiac anomalies, with respect to therapy settings which is generally adjusted by authorized physicians, a malicious modification of these settings may lead to a future inappropriate reaction, and thus harming the health status of the patient.

Cloud Based Services for Healthcare

Cloud computing can be considered as a promising paradigm for the enhancement of healthcare services. First, it allows coping with the limited storage of the traditional local healthcare servers, as it provides an infinite storage of medical data. Second, it ensures a continuous availability and sharing of e-health data that are managed by different healthcare providers. Third, it allows providing virtualized healthcare services (generally using a third-party application) independently of the geographic locations of the involved healthcare professionals. In fact, when medical data are stored in cloud, there is no need to perform clinical tasks in the same hospital where these data are generated. Fourth, it offers a set of benefits regarding data management like scalability, Quality of Service (QoS), cost effectiveness, and ease of access.

Cloud Storage as a Service (StaaS) is among the existing cloud computing types. It provides several services, including the storage of Electronic Health Records (EHRs) and Personal Health Records (PHRs) which represent the digital version of the health information of a patient, such as medical history, medication and allergies, and vital signs. The migration of the EHRs and PHRs to the cloud offers several advantages. First, it is possible to ensure an infinite storage of data. Second, thanks to the use of cloud storage, healthcare services will be able to interoperate between them to improve the quality of healthcare services through sharing information across healthcare organizations. In fact, the patients' EHR can be distributed at various locations due to the fact that a patient may be examined by many healthcare professionals of different specialties (e.g., primary care physicians, specialists, and therapists) and register with a set of different health insurance companies according to the insurances types. Third, the use of cloud storage relieves healthcare professionals and organizations from the infrastructure management tasks.

Typically, healthcare services generate and manage sensitive personal data (e.g., medical records) that have strict confidential, privacy and security concerns. In this context, the security of cloud based services for healthcare is a challenging issue that should be addressed during their development. For instance, a long term storage of healthcare data in a cloudlet belonging to a third party organization requires robust security solutions that ensure secure access controls, audit controls, authentication, and storage privacy, in order to avoid exposing theses data to unauthorized entities and security threats (e.g., disclosure, theft). Trustworthiness of cloud environment is another security issue which depends on the cloud deployment model (e.g., private, public). Private clouds seem to be trusted and more secure than public clouds, due to the fact that the security control of the physical infrastructure and the access to data is mitigated to the owner of the private cloud. Nevertheless, the owner of the private cloud needs to call for security expertise to handle the security of the cloud infrastructure, which is a heavy and expensive task. In the other side, the organization owning the public cloud infrastructure is responsible of enforcing a sufficient security policy that guarantees the protection of data in public clouds. The use of public cloud by a healthcare provider may therefore introduce additional security risks and threats, especially as security depends on the degree of trustworthiness to the provided security mechanisms in the public cloud.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A wide set of healthcare applications was proposed in the literature. These applications can be classified with respect to the services they offer to the patients. In this context, we distinguish three classes. The first, denoted by "emergency detection and response", groups the healthcare applications that are in charge ensuring a real time surveillance of specific types of parameters (e.g., physiological, movement) allowing to proactively detect emergency situations, locate patients, and provide them with the necessary health support (e.g., alert generation toward doctors which can issue remote prescriptions, or order sending ambulance to the patient's location). The second class, "disease and safety surveillance", includes healthcare applications that are designed for the surveillance of patients and elderly persons who suffer from chronic illness and need a continuous monitoring of their health status, or need to be monitored during a recovery period after a surgery or an acute event. The third class, "medication and assistance in prescription solution", groups healthcare applications, which are in charge of managing and processing the stored electronic medical records to identify the patient's disease and assist physicians in remotely delivering the appropriate healthcare prescription.

Table 1 classifies a set of references related to healthcare applications, with respect to five criteria, namely the monitored parameters, the area of surveillance, and the provided quality of service.

SECTION 2: DIGITAL CRIMES TARGETING HEALTHCARE SYSTEMS

As emerging healthcare systems rely on mobile networks and internet to provide and improve healthcare services (e.g., remote monitoring of patients, storage and sharing of medical records), they become unprotected and vulnerable to a set of security threats due to the lack of appropriate security mechanisms. A Compromised security of such type of systems could disrupt the critical functions they perform and expose the patients to privacy and life threats. Recent research works have demonstrated that healthcare systems are targeted by several types of digital crimes that can be classified according to their goals and their impact on the patients. Four classes of crimes are described hereinafter.

Table 1. Classification of healthcare applications

Application Classes	References	Monitored Parameters	The Area of Surveillance	Quality of Service
Emergency detection and response	(Ko et al., 2010)	Physiological parameters	Hospital	Optimized rate control protocols
	(Felisberto et al., 2015)	Movement Parameters	Home	-
	(Almadani et al., 2015)	Health status parameters	Ambulance	Real time transmission
Disease and safety surveillance	(Chen et al., 2014)	ECG	Home	-
	(Al-Qurishi et al., 2015).	Multimedia and medical Information	Home	-
	(Aquino-Santos et al., 2013)	ECG	-	-
Medication and assistance in prescription solution	(Kaur & Chana, 2014)	Health data	Home	Resource elasticity mechanisms
	(Maia et al., 2014)	Vital signs of patients	Home	-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Crimes Threatening the Privacy of Patients

As healthcare applications have significant security weaknesses, several attacks ranging from eavesdropping to location tracking could be executed to discover and disclose the personal information of the victim patient. Using the discovered information, a set of crimes targeting the patients' privacy may be performed. Below, we present some examples.

Eavesdropping attack can be performed by simply snooping on the radio communication and analyzing the intercepted traffic. If the latter contains sensitive information about persons, such an attack may threaten their privacy. For instance, we consider the case of WBAN based health monitoring application. When executing an eavesdropping attack on the traffic exchanged between wearable sensor nodes and the BAN gateway, an adversary would be able to discover a set of health information (in the form of values of some physiological parameters). This can be achieved through the analysis of the intercepted traffic, if the latter contain clear text data. However, when the exchanged traffic is secured using the same cryptographic keys for a long period of time, the adversary can execute some attacks in persistent mode (e.g., brute-force attacks) to discover the used encryption keys, and therefore he will be able to decrypt the intercepted traffic. One of the criminal consequences of eavesdropping attack is the disclosure of the discovered health information which may cause significant harms to the victim, especially if these data information show the presence of chronic illnesses or other types of serious diseases. For example, a person may fail physical exams for employment (e.g., a footballer) due to the diffusion of its health information showing the presence of chronic diseases. The disclosure of health information can also cause financial harms to the victim, if for instance, his/her personal health information, which exhibits a chronic illness (e.g., cardiac arrhythmia), is disclosed and forwarded to the insurance organization which provides to him the healthcare coverage. If such a situation happens, the health insurance premium cost will be increased.

Fingerprint and Timing-based Snooping attack (Srinivasan et al., 2008) is one of the side channel attacks that can also threaten the privacy of persons who use WSN based e-health system. Such an attack can be performed by simply eavesdropping and analyzing the wireless transmission of the traffic of the sensor nodes which are deployed in home to collect contextual information. Even if all wireless communication between sensor nodes and the gateway are fully encrypted, the execution of this attack remains possible. In fact, its execution requires only the identification of timestamp and the fingerprint (in the form of a set of features of a Radio Frequency waveform that are particular to unique transmitter) of each radio message. An adversary, who performs fingerprint and timing-based snooping attack, would be able to infer the private activities of a person (e.g., cooking, showering, or sleeping). Nevertheless, the tracking of the persons' private activities in his/her home constitutes a significant invasion of privacy, which can reveal many things about his/her life, and thus can prepare for several serious harms. For instance, an adversary can post the daily activities of the victim on a social site. This would make the victim in embarrassing position, especially if he is a famous person (e.g., actor or politician).

Crimes Targeting the System Availability

Crimes targeting the availability of healthcare systems are those that are intended to eliminate the system capacity of performing the expected functions, so that the threatened medical systems would not be able to react properly during emergency situations. These crimes can be typically achieved by disrupting the performance of the deployed network or by altering the functionality of the system components. Below, we detail some examples of crimes targeting the availability of e-health systems.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jamming is defined as the continuous emission of high power interference to disrupt the wireless communication between two communicating entities. When executed, this attack may disrupt the availability of the system, affecting the instantaneous response of the system, which is needed to ensure the patients safety. Typically, when executed on e-health monitoring applications, jamming attack can target the traffic exchanged between sensor nodes belonging to the same network or between the wireless network (e.g., WSN or WBAN) and the gateway to prevent the collected data from being transmitted to the gateway, and thus from being forwarded to the analysis center. Nevertheless, if the jammed traffic contains information showing the detection of an emergency situation, the system could not trigger the appropriate reaction to such situation. By the way of example, we consider a Home Care Monitoring System (Magaña-Espinoza et al., 2014) which allows monitoring the heart rate of patients in their home and triggering alerts to physicians or family members when detecting emergency situations (i.e., rapid physiological changes due to falls or occurrence of some types of fatal arrhythmias like tachycardia). By executing jamming attack on the messages exchanged between the nodes monitoring the patient and the gateway, the sensed data could not be delivered to the gateway, so that it may become impossible to detect the occurrence of emergency situations, and thus it will be impossible to generate instantaneous alerts towards the involved physicians.

Sensor worm attack (Tan et al., 2011) can also be classified among the security attacks threatening the availability of WSNs for healthcare applications. Such an attack consists in discovering and exploiting software vulnerabilities of sensor nodes. It can be executed as follows. The adversary discovers exploitable software vulnerabilities (this may be achieved through the analysis of the libraries of the popular operating systems running on sensor nodes). Then, he exploits the vulnerabilities of the implemented software update mechanisms to inject a malicious program into the firmware of the sensor node. For example, when sensor nodes implement an over-the air-programming mechanism to update the healthcare applications, the adversary can exploit the self-programming routines in the bootloader to inject a malicious program into these sensor nodes. Typically, the injected malicious codes allow tampering the logic of sensor nodes and altering the expected sensing and processing functions they implement. For instance, when performing counterfeit software updates of medical sensors belonging to healthcare WBANs, these nodes would not able to accurately measure and collect physiological parameters.

Crimes Targeting the Life of Patients

As discussed in the previous section, Implantable medical devices are autonomous medical devices. They are wirelessly accessible through an external programmer. Nevertheless, because of the use of insecure and vulnerable communication protocols, IMDs are threatened by a set of security attacks, where most of them could be a serious source of crimes targeting the life of patients carrying them. In the sequel, we detail some of these security attacks.

Battery depletion attack can exhaust the battery of IMDs by flooding these devices with a large amount of bogus encrypted traffic. In fact, for security purposes, an IMD is required to verify the authenticity of every received message using cryptographic computations before proceeding with acceptance or rejection. By receiving a huge amount of encrypted messages, the IMD consumes a lot of energy, due to the intensive execution of the implemented cryptographic algorithms. Consequently, its battery may be depleted, as it has a limited lifespan. When the battery depletion attack targets some types of IMDs, such as defibrillators, the consequences of the battery depletion attack may be fatal. In fact, the defibrillators

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

could not perform the implemented life-sustaining functions including the detection of arrhythmias and the delivery of therapies. In (Hei et al., 2010), it was shown that a depletion attack on an Implantable Cardiac Defibrillator (ICD) can be easily launched by an adversary who uses a radio software platform to conduct a reverse-engineering of the communication protocol used by the ICD and execute the attack.

Exposing some types of IMDs, such as pacemakers, to electromagnetic interference (EMI) can be also classified among fatal attacks threatening patients carrying these devices. The use of EMI within a frequency spectrum being detected by the sensing circuitry of the pacemaker, can affect the pacemaker functioning in different ways. When exposing the pacemaker to a magnetic field that is within the heart rate range, the pacemaker will stop the delivery of the stimulating pulses that regulate the heart's rhythm. Consequently, patients will be susceptible to lightheadedness, syncope, or even death according to the duration of the EMI and the cardiac rhythms. Otherwise, when the pacemaker is exposed to a strong magnetic field, it contiguously delivers stimulating pulses at a fixed rate (which can be rapid or low rate pacing) regardless the real heart's rhythm. This can threaten the safety of the patient, since a very rapid pacing can induce the occurrence of ventricular fibrillation, which is one of the fatal arrhythmia.

Another fatal security attack consists in using a relay node to repetitively replay old messages which are typically exchanged between an implanted device or sensor node, and a remote controller. The reception of old messages that contains old values of the sensed physiological parameters leads the remote controller to order the delivery of erroneous treatment (e.g., mistreatment or overtreatment) which can endanger the health status or even the life of the patient. For instance, insulin delivery systems are vulnerable to replay attack (Li et al., 2011), due to the weakness of protection against replay attack (every packet showing a counter different to the one received in the last packet is accepted). In fact, by replaying in an altering fashion two intercepted packets, an outdated glucose level can be reported to the insulin pump. Moreover, the replay of an old message to the insulin pump showing high blood glucose level will cause the delivery of an overdose of insulin.

Crimes Targeting Medical Records

Medical records of a patient may refer to Personal Health Information (PHI), Electronic Medical Record (EMR), or Electronic Health Record (EHR). Typically, these records include very sensitive data that should be properly protected against security and privacy threats, since their unauthorized alteration or disclosure can induce serious consequences to the victim ranging from financial to life threatening harms. Nevertheless, due to the complexity of the architecture of e-health systems managing these records and the lack of security solutions that ensure end to end protection of medical records (during transmission, storage, and access), medical records are threatened by a set of security attacks.

The use of unsecure end-user systems to access to medical records makes them vulnerable. End-user systems (e.g., personal computers or Smartphone) are generally used by an authorized patient or physician to retrieve or update medical records. However, access to medical records using end-user systems is sometimes performed using weakly protected operating systems, or even weak authentication techniques that are implemented by third party applications for securing access to medical records. As discussed in the work of Dmitrienko et al. (2013), an adversary may exploit the vulnerabilities of the software running on the end-user system to get unauthorized access to medical records, by inducing this victim user to install malicious software (e.g., malware) or faked third party applications. When the victim user uses a fake third party application to log, the credential he enters will be sent to the adversary.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Medical identity theft is one of the serious consequences of unauthorized access to medical records. It presents one of the vicious new kinds of crimes threatening medical records. Medical identity theft is defined as the fraudulent use of personal health information to obtain medical services or goods such as prescription drugs, medical treatment, or Medicare reimbursement. It often aims to alter or inject false medical records (e.g., false diagnosis, falsified medical history) to make healthcare professionals unable to provide the appropriate treatments. Medical identity theft can also lead to a fraudulent acquisition of medical services, inducing an important financial loss to the victim.

SECTION 3: REQUIREMENTS AND CHALLENGING ISSUES FOR A POSTMORTEM INVESTIGATION OF HEALTHCARE CRIMES

A postmortem digital investigation system of healthcare crimes should at least fulfill the following requirements.

Ubiquity and Diversity of Collected Evidence

To be investigable, a healthcare system should provide a variety of accurate ubiquitous evidence. In fact, due to the diversity of the implemented technologies and services in healthcare systems, a postmortem investigation of these systems should address the investigation of the entire system components including networks, databases, embedded sensor nodes, and medical devices. To this end, the following types of evidence should be at least available:

1. Evidential traces collected from sensor nodes. These traces could help identifying the source of the malicious nodes, and the malicious forged/exchanged information, which are used to prepare for the crime execution;
2. Evidence data provided by wireless access devices (e.g., gateways or access points). Typically, these data contain traces related to the sensitive activities (e.g., authentication, database update) performed by remote users. They can help identifying potential crimes on networked healthcare systems;
3. Digital traces collected from autonomous medical devices, capable of automatically ordering/delivering therapy (e.g., Implantable Medical Devices). This type of evidence would allow identifying inappropriate functionalities performed by these devices and representing the root of the incident damages;
4. Database evidence traces (e.g., EHR, contextual information, physiological parameters) which include the data collected by healthcare devices. They could allow identifying potential crimes targeting the life of patients, since they would be useful to determine how and when the health of the victim was deteriorated;
5. Digital evidential traces related to the third-party applications. The latter traces may be useful in some circumstances (e.g., when the involved third-party application is malicious or indirectly used to facilitate the execution of the crime), since they would allow identifying the potential attacks exploiting these applications and determining their objectives;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

6. Evidential traces provided by physical authentications devices (e.g., smart cards). In fact, since these devices are used to authenticate users, they store various information including authentication keys, and some user-generated sensitive activities. Such type of information would be useful to determine the behavior of the user carrying the physical authentication device; and
7. Evidential traces provided by virtual infrastructures. Since in some cases, a virtual infrastructure (e.g., virtual cloud instance) can be the place where the incident happened or originated, it can provide useful information for forensic investigation.

Hypothetical Investigation Tolerant to Data Incompleteness

Typically, an investigated system cannot totally detect and record all executed events. To this end, evidential data provided to the investigator may be incomplete. In some circumstances, investigators may also be faced to new types of security attacks targeting the system under investigation. Moreover, the implemented library of attacks can also exhibit the absence of some security attacks targeting the investigated system. Nevertheless, to enable the reconstruction of potential crime scenarios, the postmortem investigation system should be tolerant to the incompleteness of the provided information related to evidential traces, investigator knowledge, and the library of attacks. To do this, such a system should allow hypothesizing missing events and generating hypothetical actions when needed.

Accuracy and Correctness of the Postmortem Investigation System

Because of the diversity of the collected evidence and the complexity of crime scenarios, the results of postmortem investigation may become prone to errors. To cope with this, postmortem investigation should be based on formal techniques and methodologies. The latter are time-saving and able to automatically infer provable potential crime scenarios starting from the collected evidence, and provide accurate and relevant results, even with a huge amount of evidential data.

Trustworthy and Evidentiary Aspect of Evidence

In order to enable an efficient postmortem investigation, the trustworthy and integrity of the provided evidential traces are stringent requirements. Moreover, once gathered, digital evidence should fulfill the same legal requirements. They should be:

1. **Complete:** Contain enough information allowing proving or disproving the occurrence of the crime;
2. **Reliable:** Have not been changed after their creation and they are collected using reliable techniques from a physically protected device that ensures at least tamper proof evidence;
3. **Believable:** Convincing to juries; and
4. **Admissible:** Collected using procedures that are conform to the common law and legislative rules.

To this end, it becomes crucial to protect the evidential traces against security attacks, especially from anti-forensic attacks which are intended to alter the generated evidence to hide information related to the executed attack, and mislead the investigator. The use of tamper-evident mechanisms in e-health system could be a promising solution. In fact, such a mechanism could allow tracking and recording all actions (e.g., removal, modification) that may be performed on the digital traces, so that the integrity of these

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

traces can be checked before postmortem analysis. Moreover, it is also required to implement a security solution, which allows the protection of e-health systems against security attacks and the provision of a secure environment for the measurement, the collection, and the storage of the evidential traces.

Aggregation of the Pathologists and Healthcare Providers' Efforts, and Digital Forensic Investigators Efforts

As digital crimes on healthcare applications may affect the life of patients, a postmortem investigation of crimes on healthcare systems requires fitting together the efforts and the decisions of healthcare experts (e.g., physicians, pathologists) and digital forensic investigators. In fact, when a digital crime was conducted on a healthcare system and led to the deterioration of the health status of the victim, digital investigators can only demonstrate its occurrence and reconstruct plausible crime scenarios starting from the collected digital evidence. On the other side, healthcare experts can only prove that the inappropriate functioning of the healthcare system induced the deterioration of the patient health status. Nevertheless, none of them (i.e., healthcare experts and investigators) could prove or disprove whether the death is a consequence of a digital crime, unless the causality between them is identified. This can only be achieved by conducting a collaborative postmortem analysis between forensic investigators and healthcare experts. Nevertheless, to provide accurate results, forensic investigators and healthcare experts need to make accessible and share between them the collected evidential data and the obtained results. The development of forensic as a service cloud-based solution can be a promising for the collaborative postmortem investigation. In fact, such a service would allow benefiting from the expertise of investigators, as it makes collaborative investigation available anytime at any place. In addition, it can also help constructing a large library of investigated crimes scenarios, which can be useful to improve the efficiency and the productivity of the forensic experts. A postmortem investigation can also benefit from the storage as a service, as it allows coping with the storage limitation of the provided evidential data. In fact, such a service can provide an infinite storage space of evidential data. Moreover, it could also allow sharing evidential data across the involved investigators.

SECTION 4: AUDITING AND INVESTIGATING CRIMES ON E-HEALTH SYSTEMS

E-health systems provide convenience and efficiency to elderly and chronically ill persons through the healthcare services they implement. Nevertheless, as discussed in Section 2, healthcare systems are vulnerable to a set of digital crimes, which can threaten the privacy, the safety, and even the life of persons using these systems. To cope with this, it becomes crucial to develop forensic readiness e-health systems that enable the secure auditing and the postmortem investigation of crimes on these systems. In the context of digital crimes targeting healthcare systems, a pre-mortem investigation could not be applied because of the need of the presence of forensic pathologists to collect the medical evidence. The latter are essential for the reconstruction of medical scenarios.

The existing postmortem investigation approaches (Carrier & Spafford, 2006) tailored to general types of digital systems cannot be applied directly to investigate digital crimes on e-health systems, for the following reasons. First, contrary to the general types of digital systems, where a crime affects only the hardware of software assets, crimes on e-health systems affect these types of assets as well as the patient. The consequences of these crimes are more serious, as they affect the health status of the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

patient and may even cause its death. Second, when reconstructing the potential executed scenarios, the classical digital investigation techniques assume that the consequence of an action (malicious or legitimate) part of the scenario is deterministic. This is not the case for the investigation of several types of healthcare crimes. In fact, the consequences of an action may vary (i.e., is probabilistic) depending on the behavior, the health status, and the physiological state of the patient. Third, conducting an investigation of digital crimes on e-health systems requires the collaboration of forensic pathologists and digital investigation experts. The forensic pathologists will be in charge of examining the status of the victims and determining the cause of death and/or injuries, while the forensic digital investigation experts will be in charge of examining digital evidence and reconstructing the scenarios of attacks (e.g., illegal modification of sensors configuration), determining the source of the attack, and identifying the types of exploited vulnerabilities. Proving that an illegal modification of the e-health system is the root cause of an unsuitable medication that leads to the death cannot be done unless the conclusions derived by the two types of experts are correlated together. Fourth, a healthcare attack could take a long time to impact the patient. For example, an illegal modification of the electronic healthcare records a patient may lead to an unsuitable medical several months or years after. Reconstructing the whole scenarios and finding the traces related to the illegal modification, requires storing the evidential data for a very long time. Such a feature, introduces additional challenges related to the long-term secure storage of evidence.

Several research works were proposed in the literature to address the secure auditing and the postmortem investigation of digital crimes on e-health systems. In the sequel, we provide an overview the most important ones of these research works.

Auditing the Health Information Exchange

Health Information Exchanges (HIEs) systems refer to distributed systems ensuring the sharing of medical records between healthcare institutions. As discussed in Section 2, the inappropriate sharing of medical records, which are highly sensitive data, could threaten the privacy and the safety of patients. It is therefore essential to implement techniques and mechanisms, which allow providing secure accountability and auditing of the exchanges of these sensitive data, especially as evidential traces (i.e., audit data) could be useful to conduct postmortem investigation of crimes on e-health systems. These techniques will be in charge of monitoring the exchanges and the sharing of health information, and logging sensitive action executed on e-healthcare systems, including access, share, and transfer of healthcare records. To allow tracing the entire sharing path of a patient's health records, the logged data should contain a wide set of information, including, but are not limited to, the identities of the entities (e.g., the physician, the hospital) involved in the creation of the health record, the timestamp of these records, and the identities of the entities updating them.

To improve the accountability of actions executed in e-health systems and enable the identification of data breach, Mashima and Ahamad (2012) developed a patient-centric scheme allowing the establishment of robust information accountability in electronic health record (EHR) sharing systems. The proposed system introduces the use of a secure metadata, namely "accountability tag", which is attached to each copy of EHR record. An accountability tag implements a cryptographically-verifiable evidence of the involved entities (e.g., the repository that creates the copy of the healthcare record, the sharing source, and the sharing destination). Thanks to the use of accountability tags, the proposed scheme offers the benefit of patients' awareness to the access and the sharing of EHR in the e-healthcare system, so that a patient can determine whether sharing is appropriate and even identifies the involved entities when

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

detecting a suspicious disclosure or sharing of its healthcare data. Accountability tags are managed by a trusted agent, namely "patient-centric monitoring agent". They are verified and logged each time the "Accountable Update" or the "Accountable Usage" protocol is executed, so that the monitoring agent will be able to reconstruct the sharing path and identify the involved entities. Accountable Update protocol is initiated when a health record is created or updated. In fact, when the record issuer creates or updates a health record, he attaches his digital signature, which is encrypted using the public key of the monitoring agent, to the created record. After that, he sends them to the repository which will be in charge of obtaining the proof of patient's authorization through the monitoring agent. After accepting the creation or the update of the record, the repository generates and sends verifiable receipts to the monitoring agent and the record issuer. The Accountable Usage protocol is executed when the consumer receives the healthcare record. Since the consumer needs to decrypt the received signature, he sends a verification request to the monitoring agent. The latter logs this action, decrypts the signature, and creates a non transitive signature using Universal Designated Verifier Signature scheme (Steinfeld et al., 2003). Such a signature can be only verified by the specific consumer. After that, the monitoring agent sends back the generated signature along with a transaction proof to the consumer which verifies this signature using the issuer's public key.

For the sake of privacy preservation, Oh et al. (2014) proposed a Health Information Exchange (HIE) audit system which allows limiting the amount of accessible logs' information to only those that are required to provide evidence for audits process. This system implements together the use of Hierarchical Identity Based Encryption (HIBE) algorithm, which is an encryption algorithm ensuring the encryption of the data stored on audit logs, and a logic-based audit algorithm, which is able to handle the incompleteness of the available logs and generate formal explanations using only the available decrypted information. The proposed audit system uses two types of logs: internal Audit Trail and Node Authentication (ATNA) logs, which represent standard audit logs providing infrastructure for HIE audits, and external audit logs which include supplementary information (e.g., patient registration, medical billing) retrieved from healthcare organizations to enhance the results of HIE audits. The audit system mainly integrates four components: the audit data processor (ADP), the audit agent, the audit algorithm, and the auditor viewer. The ADP is in charge of encrypting external audit logs using HIBE algorithms, aggregating ATNA logs and external logs together, and making them available to the audit agent. When receiving a set of accesses of audit interest from the auditor through the auditor viewer, the audit agent transmits audit logs together with the policy and the requested access of interest to the audit algorithm, which will be in charge of verifying whether the requested access of interest complies with the policy. Since audit logs are encrypted, the audit algorithm may not be able to decide about compliance or violation. To this end, starting from the residual policy, the audit agent identifies the minimum data information required for audits, requests the corresponding decryption keys from the ADP, decrypts these data information, and then makes them available to the audit algorithm. Based on the available decrypted data, the audit algorithm generates the formal explanations for accesses and then sends them to the auditor viewer, which will be in charge of translating these explanations in human readable information. After that, it returns them to the viewer. When detecting incompliance, the auditor sends a notification to the involved healthcare organizations.

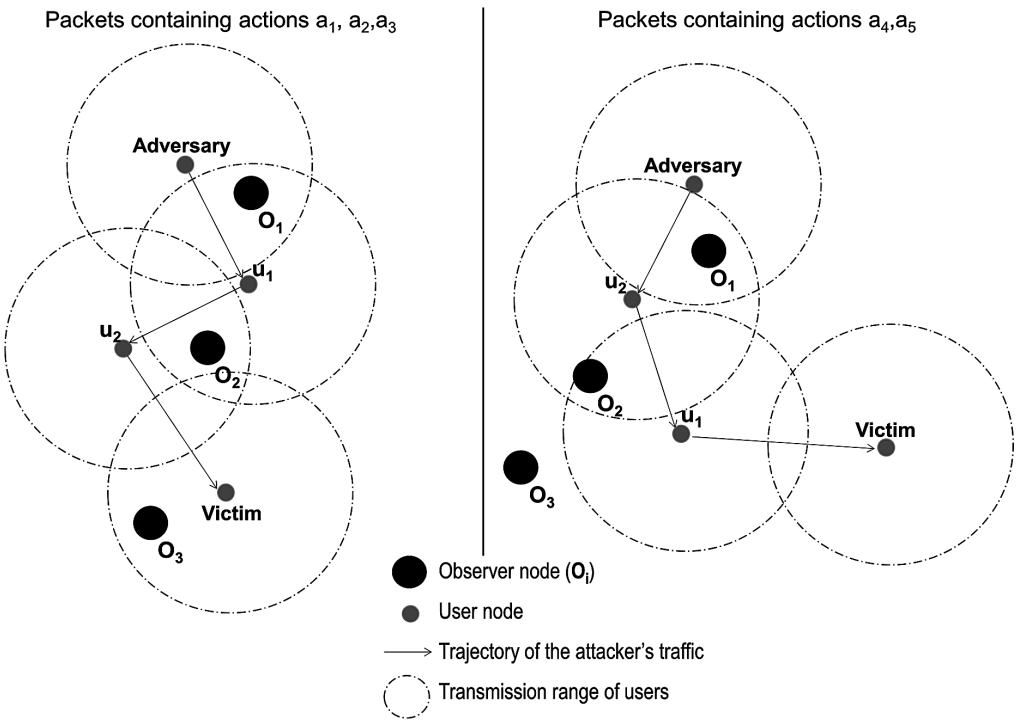
EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Forensics in Wireless Medical Body Area Networks

As detailed in Section 2, wireless medical Body Area Networks (BANs) are vulnerable to specific classes of security attacks that can be exploited to perform crimes threatening the privacy and the safety of patients who use them. In this context, it becomes crucial to address forensic investigation of security attacks on wireless medical BANs. Several research works investigated the possibility of using wireless medical BANs for healthcare applications. Nevertheless, to the best of our knowledge, there is no postmortem investigation system proposed in the literature, which is tailored to wireless Medical BANs. The work of Cusack and Kyaw (2012) discussed only the need of developing a forensically ready and secure wireless Medical BANs based systems, which allows investigating events after their occurrence. However, it is possible to customize the forensic systems, proposed in the literature to investigate security attacks on Mobile Ad hoc and Sensor Networks (MASNets), to the investigation of healthcare security attacks and crimes in wireless medical BANs, especially as wireless medical BANs are specific classes of MASNets.

Since Mobile Ad hoc Sensor Networks (MASNets) are typically characterized by the absence of infrastructures and the mobility of nodes, techniques and mechanisms developed for the investigation of attacks on wired networks cannot be efficient to investigate attacks on MANets. In fact, as attacks targeting such type of networks are generally mobile, traditional evidence collection techniques for wired networks could not allow tracing and tracking mobile attacks. To cope with this, a digital investigation in MASNets needs to implement new techniques allowing a distributed supervision of the network, so that it becomes possible to collect digital evidence from different locations in the network and the victim system. In this context, Rekhis and Boudriga (2011) developed a new architecture and evidence collection

Figure 2. Variation of the network topology during an attack scenario



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

technique in MANets which ensures an efficient collection of mobility-based information. In addition to the user nodes (i.e., mobile devices of patients) which are randomly distributed over the network, the proposed technique integrates the use of a set of mobile trusted nodes, namely observers, which are responsible of supervising the mobile networks. As depicted in Figure 2, observer nodes, denoted by $\{O_i, i \in [1,n]\}$ where is the number of observers, are distributed over the network in a way that they form an observation network allowing the supervision of every pair of communicating user nodes, denoted by u_i. Each deployed observer is responsible of monitoring the traffic exchanged within its transmission range, storing the observable parts of the detected datagrams and actions as network evidential traces, and forwarding them to the node in charge of investigation (selected among the observer nodes).

Because of the mobility of observer nodes and the scheduling algorithm they implement to minimize energy consumption, an observer node will not be able to detect all the datagrams sent by the attacker. In fact, the observer node may go out of the attacker transmission range or change its state to inactive, so that some of the attacker datagrams will be invisible. Nevertheless, datagrams, which are invisible to a given observer, can be visible to another, and thus observer nodes can generate redundant or/and complementary evidential traces which need to be merged and aggregated by the node in charge of investigation before reconstructing the occurred scenario. An example of nodes mobility and topology variation during an attack scenario is depicted in Figure 2. During the occurrence of the attack scenario, user nodes as well as observer nodes change their positions, where two paths are used to forward the datagrams sent by the adversary to the victim node. While the first path (Adversary $\rightarrow u_1 \rightarrow u_2 \rightarrow$ victim) is used to forward packets containing actions a_1, a_2, and a_3, the second path (Adversary $\rightarrow u_2 \rightarrow u_1 \rightarrow$ victim) is used to forward packets containing actions a_4, and a_5. Observer nodes take two different positions during the execution of the attack scenario. In the first position, O_1 is under the coverage of adversary and u_1 nodes, O_2 is under the coverage of u_1 and u_2 nodes, and O_3 is under the coverage of the victim node. In this position, all observers O_1, O_2, and O_3 detect the datagrams containing the actions a_1, a_2, and a_3. In the second positions, while O_1 and O_2 detect the datagrams containing the actions a_4, and a_5, since they are under the transmission coverage of adversary and u_2, and u_2 and u_1, respectively, O_3 cannot detect any packet.

Network evidence could not provide sufficient information related to some type of attack scenarios targeting MANets. In fact, attack scenarios may include some suspicious events compromising the target system, which are unobservable at the network level (e.g., a set of actions are executed locally on the target system, in response to a remote actions executed over the network) or invisible to the developed network security solution used the observers (e.g., actions are encrypted). To cope with this, Rekhis and Boudriga (2011 proposed to use and aggregate system evidence, in addition to network evidence, for conducting a postmortem investigation on MASNets. Note that, sstem evidences represent the evidential digital traces generated by the security solution that the system implements. They provide a set of information related to the observable parts of executed actions and a description of the related consequent system states.

The huge amounts of evidence that may be generated by the implemented security and network solutions can make their analysis and the generation of attack scenarios complex and prone to errors. Such a limitation can be handled through the development of a formal methodology managing evidence and ensuring an automatic reconstruction of possible attack scenario. In this context, the methodology for investigation of security attacks proposed in the work of Rekhis and Boudriga (2011) implements an inference system that integrates network and system evidence and copes with the incompleteness and duplication of the information to generate the possible and provable attack scenarios. The proposed methodology is organized into four principal steps described as follows. During the first step, the node

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in charge of investigation will be in charge of collecting the available evidence from the victim system and the observer nodes deployed in the network under investigation. The network evidence should be also filtered to include only the digital traces related to the attack under investigation. In the second step, an aggregation of the collected network evidence is performed. This can be achieved through the execution of a set of inference rules allowing the generation of merged and optimal network evidence. Heuristics, which are based on information appended by nodes during the generation of their observations, are also implemented to retain plausible sequences. The third step consists in the generation of the set of possible attack scenarios based on the provided network and system evidence. The fourth step consists in determining the provable actions and scenarios among the generated possible attack scenarios. All of the previous detailed steps are automatically performed through the execution of a specific set of inference rules, which allows aggregating the generated network evidence, inferring possible attack scenarios, and generating provable actions and scenarios.

In the case when the fourth step fails to determine a provable attack scenario, investigators are expected to select from the potential evidence a plausible attack scenario. Three techniques can be used. In the first technique, the set of information, which profiles the attacker, is extracted from every possible attack scenario. Based on the collected information, the scenario showing the most malicious profile will be selected as the plausible attack scenario. In the second technique, the set of information, which indicates the estimated degree of damage of every attacked system resource, is extracted from every generated possible attack scenario. The scenario showing the highest degree of system damage will be considered as the most plausible attack scenario. In the third technique, every action in a possible attack scenario is labeled by the probability value of its occurrence from the state in which it is enabled. According to the probability values, the plausible attack scenario will be selected.

The inference conducted in this research is a sort of data mining by which the investigator can access to a knowledge base describing known attacks on systems, and the description of the investigated system, detect interesting dependencies in a very noisy data, and infer more elaborated scenarios involving the execution of several actions that argument the content of the collected evidence. The use of data mining for the purpose of investigation is developed in other chapters of the same book.

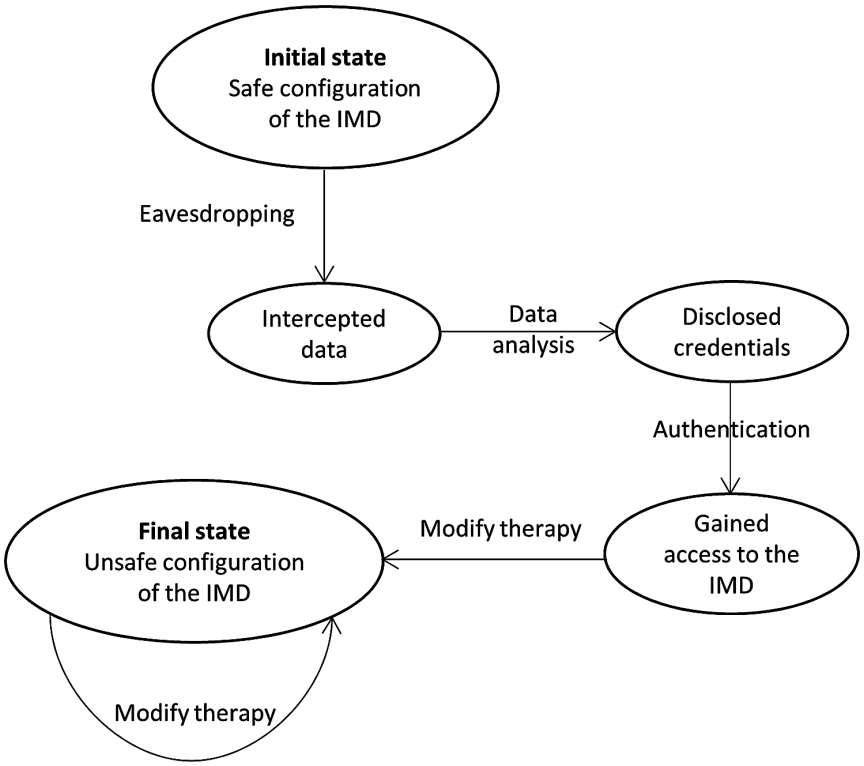
Digital Investigation of Criminal Attacks on Implantable Medical Devices

As discussed in Section 2, IMDs are vulnerable to a set of attacks threatening the life of patients carrying them. Some of these attack scenarios are illustrated in the graph depicted in Figure 3. Every node in the graph specifies the state of the IMD, which can be secure (safe configuration) or unsecure (safe configuration). Every arrow in this graph represents an elementary attack action (e.g., eavesdropping, modify therapy, data analysis). An example of a lethal attack scenario depicted in this figure is described as follows. An attacker intercepts the traffic exchanged between the IMD and an authorized programmer (A_1: eavesdropping), and analyzes the intercepted data to disclose the credentials used for authentication (A_2: data analysis). Having identified the credentials, the attacker authenticates himself to the IMD and gets a privileged access to the IMD (A_3: authentication). After that, he modifies the therapy settings (A_4: modify therapy), so that the IMD configuration becomes unsafe. A further reaction of the IMD in the case of heart arrhythmia, would be inappropriate, and could harm the health of the patient.

In this context, the design of a postmortem investigation system of lethal attacks on cardiac IMDs becomes crucial. This system would enable the identification and the reconstruction of potential attack scenarios source of death of the victim. Such a system should combine together medical and technical

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Examples of lethal attack scenarios targeting IMDs



notices. While the technical investigator conducts forensic analysis to prove or disprove the occurrence of malicious attacks, the medical expert performs forensic analysis to prove that the IMD malfunction is the source of death of the victim.

Typically, cardiac IMDs (e.g., pacemakers and cardiac defibrillators) provide a set of digital traces such as, the history of electromyogram and the history of the delivered therapy. These traces could be useful to conduct a forensic investigation. In fact, starting from these traces, pathologists can derive some conclusions about the primary cause of death. Nevertheless, these traces are insufficient to provide accurate investigation results. They need to be extended to include the sensitive events that may be executed on the IMD. Moreover, they should be authenticated and secured against tampering and forging. In addition, since a cardiac IMD may provide a huge amount of evidential traces, it becomes necessary to develop appropriate mathematical and formal analysis techniques for postmortem investigation allowing the automatic generation of results (i.e., potential scenarios). In this context, Ellouze et al. (2014) proposed a postmortem investigation system of lethal security incidents on cardiac IMDs. Such a system allows the generation of potential attack scenarios source of the death of the patient carrying a cardiac IMD. It reconciles in the same framework the conclusions derived by forensic medical experts and technical investigators.

The proposed system defines two types of evidence: medical and technical. Medical evidence refers to the medical experts (e.g., pathologists) findings during a forensic analysis of the victim's body and IMD. Typically, medical evidence include:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. The history of the electromyogram (EMG) of the patient over the period of time preceding the death;
2. The set of therapies (i.e., electric shocks) delivered by the IMD, as a series of time stamped reactions along with the value of electric shocks;
3. The clinical observations which are derived during autopsy; and
4. The information collected from the medical record of the victim.

Technical evidence present the set of sensitive events executed on the IMD. They are typically stored in IMD logs as a series of time stamped events. They include:

1. Access information which help differentiating between sessions of remote connection;
2. Evidential traces related to the updates of IMD configuration and therapy settings;
3. Evidential data showing the occurred sensitive actions (e.g., the update of the IMD firmware); and
4. Information of the traffic exchanged between the IMD and remote devices. Such information is useful for the identification of some security attacks (e.g., battery depletion).

As the results of postmortem investigation mainly rely on the correctness and the trustworthy of the provided evidence, Ellouze et al. (2014) proposed an in depth security solution, which allows securing cardiac IMDs against security attacks and protecting the provided evidence from tampering to ensure their integrity. The proposed solution implements two security layers. The first layer is responsible for the protection of the cardiac IMD against security attacks (e.g., unauthorized access). It implements a mutual authentication protocol together with a security protocol ensuring a secure encrypted exchange between the IMD and the authorized programmer. The second layer is responsible of the protection of IMD logs against tampering and removal. It can only be executed by remote users who have been successfully authenticated in the first layer. The second layer implements also a biometric authentication protocol (Ellouze et al., 2013) allowing the verification of the physical proximity of the remote user. In fact, only authorized physicians, who are able to provide the biometric key computed over the electrocardiogram collected from the patient body, can get access to modify or delete records in IMD logs.

The proposed postmortem investigation system implements a three-step methodology. The first step consists in the medical investigation, which is responsible of the automatic generation of potential medical scenarios source of the death of the victim based on an inference engine executed in backward chaining starting from the victim death. This step is conducted based on the medical evidence collected from the cardiac IMD under investigation and a library of inference rules, which describe the causal relations between the different types of arrhythmias and the reactions undertaken by the cardiac IMD. Typically, an inference rule of this library takes the following form to state that the successive occurrence of *Ev* for n times induces the successive occurrence of *Ev* m times within a period of time T, where n and m are positive non-zero integers, and *Ev* and *Ev*' are a conjunction and/or disjunction of simple events (e.g., *Ev* is conjunction of two simple events: an arrhythmias and an instantaneous reaction undertaken by the IMD; and *Ev*' is a subsequent arrhythmia). Events included in inference rules can be classified into: a) observable events, which are the events detected by the IMD and included in the provided medical evidence. Examples of observable events include the delivery of electric shocks and the occurrence of ventricular fibrillation; and b) unobservable events, which are the events that cannot be detected by the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

IMD, so that they are not included in the medical evidence. Starting from the last observed event in the medical evidence (i.e., death), inference rules are executed in backward chaining to reconstruct a tree of potential medical scenarios, whose root node is the death event.

In the second step, which refers to the technical investigation, investigators are in charge of reconstructing the potential attacks scenarios. Starting from the library of attacks, they combine elementary actions to construct attack scenarios that provide the same collected evidence when executed. Typically, an attack scenario, say S, takes the following form $S = \langle s_0, A_1, s_1, \ldots, A_n, s_n \rangle$, where s_0 is the initial system state, $A_i (i \in [1,n])$ are the elementary actions of the scenarios, which can be legitimate or malicious, and $s_i (i \in [0,n])$ represent the description of the intermediate system states. An intermediate system state s_i is generated unless an action A_i is enabled at state s_{i-1} and is executed from it. With respect to the monitoring and security supervision mechanism implemented on the IMD, an action or a state can be visible (i.e., its value is included in the available evidence) or invisible (i.e., none information regarding its occurrence can be determined from the evidence). To support the automatic reconstruction of potential attack scenarios, a Model Checking based algorithm is executed in forward chaining. Such an algorithm is in charge of executing possible actions from the library, to generate a graph of technical scenarios satisfying the provided evidence.

After the generation of potential medical scenarios and potential technical scenarios, a correlation between these two types of scenarios is performed during the third step. This correlation allows proving whether the malicious actions in the technical scenarios of attacks caused the absence of reactions or the execution of inappropriate reactions identified in the medical scenarios. Correlation of the generated scenarios is performed in three sub-steps. The first sub-step consists in the deeply analysis of the potential medical scenario to check whether there are suspicious IMD reactions (e.g., delivery of unsuitable therapy) in the potential medical scenarios, and identify their related parameters. In the second sub-step, malicious actions threatening the IMD security are identified through the examination of the generated potential technical scenarios. A correlation is conducted in the third sub-step to check whether there is a causal relation between one of the identified suspicious IMD reactions and the determined malicious actions.

An example of a lethal attack on cardiac IMD is described as follows. An attacker, who gained an unauthorized access to the IMD, altered the thresholds associated to the detection of some arrhythmias, so that the IMD becomes unable to appropriately detect and treat the occurred arrhythmias. This inadequate configuration of the IMD has led to the occurrence of fatal arrhythmias causing the victim death. To investigate this lethal attack, the three-step methodology is performed as follows. During the first step, pathologists generate the medical scenario leading to the victim death. For instance, starting from the available EMG history and the IMD response traces, pathologists prove that the occurrence of successive events of a fatal arrhythmia (e.g., ventricular fibrillation) has induced the death of the victim as follows. They demonstrate that:

1. The IMD inappropriately detected the occurred arrhythmia events, for example it detected Sinus-Tachycardia (ST) events as Ventricular Fibrillation (VF) events. To this end, it delivered the therapy suitable for VF rather than ST;
2. The delivery of successive inappropriate therapies have led to the occurrence of VF events whose intensity grew from an event to another; and
3. The IMD was unable to handle the occurrence of VF events and deliver the suitable therapy, since it reached the maximum number of electric shocks it can deliver.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In the second step, starting from the available technical evidence and the library, which contains simple attacks targeting IMDs, technical investigators reconstruct attack scenarios, which induced the inappropriate detection of arrhythmias. For instance, they prove that the inappropriate detection of the IMD occurred further a remote attack executed on the IMD, which altered the thresholds associated to the detection of some arrhythmias (e.g., Sinus-Tachycardia). In the third step, the correlation of two types of generated scenarios (medical and technical) showed that the alteration of IMD settings (i.e., thresholds associated to the detection of some arrhythmias) makes the IMD unable to deliver the therapy suitable to the occurred arrhythmias, which led to the victim death.

SECTION 5: COMPUTER AIDED PHYSICAL CRIMES INVESTIGATION

A set of research works were proposed in the literature took interest to the development of theoretical approaches for the investigation of physical health crime scenes. In the sequel, we detail the most important ones of these approaches.

Modeling Crime Scenarios Using Assumption Based Truth Maintenance System (ATMS)

In order to improve the criminal investigation process and assist investigators, a decision support system (DSS) was proposed in the work of Keppens and Schafer (2006). The proposed system implements a knowledge driven methodology together with an abductive diagnosis approach. The principal component of the proposed architecture is an Assumption based Truth Maintenance System (ATMS), which is an inference mechanism allowing the reconstruction of all possible scenarios that can provide a set of explanations for the available evidence and support some hypotheses. Possible scenarios are generated starting from an initial set of inputs (facts and evidence) using a knowledge base. The constructed possible scenarios are then grouped in the form of a data structure, namely the scenario space. In order to analyze the scenario space, users can generate a set of standard queries which are translated into a specification of ATMS nodes of interest through the query analyzer. Answers related to the received queries are computed through the extraction of the relevant parts from the scenario space, transformed in a user understandable format by the report generator, and then sent back to users.

According to the proposed DSS, a scenario consists of a set of events, states, and causal relations between these events and states. It is constructed using a knowledge base, which integrates the use of three principal constructs. The first construct is the property definitions, which allow describing the involved types of events and states. The second construct is a set of causal rules, called scenario fragments. An example of a causal rule is as follows.

```
if {
suffers(P, C),
cause-of-death(C, P),
medical-examiner(E)
} assuming {
determine(E,cause-of-death(P)),
correct-diagnosis(E,cause-of-death(P))
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

```
} then {
cod-report(E, P, C)
}
```

This rule states that: a) if there is a person *P* who suffers from ailment or injury (denoted by *C)*, *C* is the cause of death of the person *P*, and there exists a medical examiner, denoted by *E*; and b) assuming that *E* determines the cause of death of *P* and performs the correct diagnosis; then a piece of evidence is generated in the form of a report which indicates that according to *E*, *C* is the cause of death of *P*.

The third construct consists of a set of inconsistencies, which allows identifying the combination of events, and states that are impossible and inconsistent with one another. An example of inconsistency is as follows.

```
inconsistent {
suicide-action(hanging, P),
is-hanged(P)}
```

This inconsistency states that a person P cannot kill himself (i.e., suicide) and be hanged by another person in the same time.

The proposed decision support system (Keppens and Schafer, 2006) implements two types of inference mechanisms. The first type is a backward chaining inference mechanism. It infers plausible scenarios, which explains the available evidence, starting from the available evidence and until reaching the initial state of the system. The second type is a forward chaining inference mechanism. It is executed starting from the initial system state and until reaching the available evidence. It allows the generation of new plausible scenarios, in addition to those created during the backward chaining phase.

Modeling Crime Scenarios in a Bayesian Network (BN)

A Bayesian network (BN) is a probabilistic graphical model that encodes probabilistic relationships among a collection of relevant variables. For instance, in the medical field, BN may allow representing the relationships between the occurred cardiac arrhythmias and the death of a patient. It can be also used to compute the probabilities of the occurrence of the death event. In this context, the use of a Bayesian network (BN) can be a promising solution for the modeling plausible crime scenarios and the proof of the communication between investigators and judges in legal cases (e.g., crime on e-health system). Nevertheless, the identification of relevant variables, during the construction of such a network, is not always easily feasible. To this end, a formal foundation for evidential reasoning, which allows the automatic construction of a Bayesian network, can be used to enable the development of a software tool, which allows assisting investigators and judges during the investigation of crimes on healthcare systems.

Vlek et al. (2013) proposed a formal evidential reasoning approach, which allows modeling plausible crime scenarios for a legal case in a single Bayesian Network and connecting them to the provided evidence. The proposed approach integrates four main steps described as follows. The first step consists in formulating all the relevant scenarios and identifying the guilt hypotheses that could be supported by each specified scenario. It may exist relevant scenarios which do not support any guilt hypothesizes. Guilt hypothesis should be formulated in such a way they are either equal, or mutually exclusive. During the second step, a set of scenarios idiom are generated and represented in the graph under construction

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

starting from the relevant scenarios (formulated in the first step). This can be achieved as follows. For every state or event in the formulated relevant scenario, a binary node is included in the graph under construction. Then, an arrow is drawn between dependent events or states, since the absence of an arrow between two nodes, representing events or/and states, means that they are independent from each other. After that, a set of binary scenario nodes are also included to represent every scenario formulated in the first step, and an arrow is drawn from each scenario node to all state and event nodes in that scenario. To represent the supported guilt hypothesis, a set of binary nodes are added, and a set of arrows are drawn from every scenario node to the guilt hypothesis it supports. In the third step, in order to merge the generated idiom scenarios in one Bayesian Network, the same guilt hypothesis nodes are replaced by one node, which will be connected to the different scenarios supporting this guilt hypothesis. After merging the scenarios idioms, a verification of the existence of other dependencies between nodes should be performed, in order to be modeled with arrows. Moreover, if there are nodes describing equal states, events or evidence, they are replaced by one node. Nevertheless, if there are nodes, in two different scenarios, describing conflicting states, events or evidence that cannot occur simultaneously, a constraint node is included to present the identified constraint. Finally, the fourth step consists in adding the relevant evidential nodes to the Bayesian network under construction. A binary evidential node representing each piece of evidence, collected during the investigation process, is included. These included evidential nodes are also connected to the event or state node in the scenario it supports. The resulting graph (i.e., the Bayesian network) will be connected, representing all idiom scenarios of a legal case.

Automated Inference of Past Criminal Action Instances in Digital Investigations

Because of the technological evolution, and the intensive use of digital equipment by persons to access to diverse types of services, the amount of digital evidence that should be handled during postmortem investigation is continuing to increase. In this context, James and Gladyshev (2015) proposed a signature analysis model allowing the automatic provision of data to knowledge acquisition for the involved digital investigators. The proposed model ensures the automated approximation and reconstruction of past action instances within a compromised digital system. It integrates the use of signature-based methods to provide a formal description of the different categories of signatures. In this work, a signature is expressed as follows $S=\{T_i,\theta,cm\}$, where: T_i represents the set of objects time stamps associated to an action instance, θ is the object update threshold related to the action instance, and cm defines the update consistency checking function which is specific to the category of object update patterns. Three principal time stamp-related update patterns are defined in the work of James and Gladyshev (2015). Below, we detail each pattern.

The first time stamp-update pattern, namely Core object update, defines the update pattern of Core object time stamps as follows. Core object time stamps represent the set of time stamps, denoted by Score, which are always updated to be equal to the value of the clock of the system each time a specific action is executed. This definition means that when observing a trace in a Core set, it can be noticed that an action instance might have occurred. In fact, an artifact observed in a Core set can be related to only single action. Moreover, due to the fact that time stamps belonging to Score are always updated, each time stamp should be within a certain time range of each other that is identified according to the object update threshold of the action, denoted by θ. Let us consider the example of a configuration file of a Program X that is modified each time the Program X is closed, thus the modification time stamp of this file is a Core trace for the action "Close Program X". Starting from the definition of the first pattern,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

an object time stamp update consistency function, namely CoreTest, was developed. Such a function is in charge of checking whether each object update conforms to the Core signature category based on the object update threshold and the TraceStates array that includes the values of time stamps of all objects specified in the signature. According to this function, if the difference between the recent timestamp value and the oldest time stamp value of the TraceStates array is lower than the object update threshold, the Core traces are considered as consistent, otherwise, they are inconsistent.

The second time stamp-update pattern, namely Supporting object update, defines the update pattern of Supporting object time stamps as follows. The supporting object time stamps represent the set of time stamps, denoted by $S_{support}$, which are irregularly (i.e., not always) updated each time a single particular action instance is executed, and they can only be updated by a single specific action. Similarly, to Core signatures, the detection of a trace in a supporting signature means that an action instance must have occurred, since an artifact observed in a supporting signature can be related to a single action. For example, an irregular update of time stamp may be figured when a file is cached in memory after the first execution of the first action instance. In fact, it may happen that the subsequent action instances accessed the file data cached in memory rather than the file disk, so that the trace related to this update would be unobservable on the disk, since the original file's meta-data on disk was not updated. Starting from the definition of the second pattern, an object time stamp update consistency function, called SupportTest, was derived. Such a function is in charge of testing whether each object update conforms to the supporting signature category based on the object update threshold θ and the TraceStates array that includes the values of time stamps of all objects specified in the signature. If an object update conforms to the supporting update category, and if each trace has been updated within the threshold θ, the action execution time can be defined to be approximately equal or shortly lower than the oldest time stamp in the TraceStates array. In the case where there exists a time stamp updated outside of θ from another related object time stamp, then it can be noticed that a separate instance of the same action must have happened.

The third time stamp-update pattern, namely Shared object update, defines the update pattern of Shared object time stamps as follows. Shared object time stamps represent the set of time stamps, denoted by Shared, which may be always or irregularly updated on the occurrence of each execution of multiple actions. Differently to the Core signatures and the Supporting signatures, more than one action can be associated to a trace in a Shared signature, so that the only information that can be derived from the observation of a trace in a Shared signature is that at least one of the actions associated to this trace must be executed. The access time stamp of a .dll file (.dll represents the file extension of a Dynamic-link library) can be considered as an example of a Shared object time stamp, since more than one action may induce the .dll file to be accessed. The conclusion that may be derived when only analyzing this file (without any additional information) is that each action, which is able to induce the .dll file to be accessed, has the same occurrence likelihood. An object time stamp update consistency function, called SharedTest, was derived from the definition of the Shared object update pattern. This function allows testing whether each object update conforms to the Shared object update category, and determining the actions associated to the trace. If an object update conforms to the Shared update category, and if each trace has been updated within the threshold θ, then the action execution time can be defined to be approximately equal or shortly lower than the oldest time stamp in the TraceStates array. In the case when there exists a time stamp updated outside θ from another object time stamp, then it can be noticed that a separate instance of the same action must have happened, because objects are irregularly updated. Since multiple actions can be associated to a trace, additional information are also required to enable the identification of the action causing the trace.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

SECTION 6: ADVANCED ISSUES

Several approaches were developed in the literature to conduct a digital investigation on compromised digital system. Several issues were addressed by these ranging from the need to generate provable results, address scalability, and provide resilience to anti-forensic attack. However, the whole majority of these approaches are not directly applicable to investigate healthcare attacks. In fact, contrary to general types of digital attacks which affect only the target system, crimes on e-health systems could firstly affect the medical devices, and then spread to the patient. A healthcare investigation would therefore require the cooperation between at least two types of experts: forensic pathologists and digital security investigators. Moreover, a healthcare attack can be conducted with a long-term interest (a modification of electronic healthcare records could lead to an unsuitable medication that will be prescribed a long time after), Techniques for the long-term secure storage of evidence need to be developed. As a medical attack on a healthcare system can affect people, its impact will vary from a patient to another, introducing an uncertainty in the attack consequence. Such an uncertainty needs to be handled by the existing investigation techniques.

Despite some advances in the development of frameworks for the forensic investigation of digital crimes in healthcare applications, several issues remain unsolved.

The first issue is related to the development of forensic readiness. We have shown in the previous sections of this chapter the importance of the risk associated to attacks on wireless medical systems, and their consequence on the patients and practitioners' safety, privacy, and legal liability. To make attacks on medical systems investigable, the complete healthcare system should be made forensically ready so that a complete and successful investigation can be conducted, or a root cause analysis can be performed when the security incident occurs. Developing the forensic readiness aspects in healthcare requires setting technical and non-technical measures to maximize the ability of collecting and securing evidences. These measures include, but are not limited to:

1. The collection of evidence legally while minimizing at the maximal possible any interference with the business process of the healthcare system;
2. The definition of the healthcare system usage scenarios that require a collection of evidence;
3. The identification of the potential sources of digital evidence, including the different wireless access networks used by patients and healthcare professionals, the various used cloud infrastructures and services, and the diverse access equipment;
4. The maximization of the capabilities of collecting evidence from the healthcare system in a way that keeps them authentic, secure, and admissible in courts of laws;
5. Training the technical staff to conduct future digital investigation; and
6. Developing the procedures to facilitate the postmortem analysis.

All of these mechanisms should be made possible, and capable to cope with the voluminous aspects of the collected medical data, the huge number of intervening persons, the diversity of the access equipment, the distributed aspects of healthcare systems and applications, the heterogeneity of the evidence to be collected, and the complexity of healthcare investigation.

The second aspect is related to the need to maximize the ability of a healthcare crime investigation system to cope with anti-forensic attacks. These attacks target the availability and usefulness of digital evidences to the healthcare forensic investigation process. The evidence can be compromised by hiding,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

destroying, or forging them. An attacker can conduct a successful crime and later alter the generated traces to make it non investigable. Aware of these types of attacks, the developed frameworks for the investigation of healthcare crimes should develop new techniques to first detect anti-forensic attacks, and also to conceal their effect on the compromised evidence. Once the anti-forensic attacks are identified, it becomes possible to proceed with the scenarios reconstruction, and the crime identification. Several issues make these actions challenging. First, the anti-forensic attacks are difficult to detect as most of them do not leave traces on their execution. Wiping an evidence file to completely hide the original content, for example, is not only difficult to detect, but also impossible to recover. Second, several classes of ant-forensic attacks exist (e.g., address anonymization, multimedia steganography, data hiding in slack space), making it necessary to develop the required formalism to model them within libraries of attacks, useable during an investigation. Third, the protection from anti-forensic attacks requires the development of novel tamper-evident technologies to make unauthorized access to the device storing the evidence easily detected. Additional levels of protection are also required to identify the potential tampers and their gained level of knowledge, determine the techniques and methods used to attempt to gain access, control access to the evidence, and reduce to the maximum possible the cost associated to the tamper evident features and the detection of tampering.

The third issue is related to the need to cope with the resources scarcity of the wireless healthcare systems, especially as they are showing an extensive use of small and tiny devices for the purpose of sensing, monitoring, identification, and on-demand information exchange. While several approaches attempted to overcome the limitation of resources and the transient unavailability of the collection devices, by using virtualization and cloud computing technologies, an extensive work still needs to be developed. The aim is not only to virtualize the storage space and the processing resources, but also to virtualize all the types of sensors used by the healthcare solution, in order to decrease the number of required sensors, increase the accuracy of the sensed data by taking profit from the redundancy of sensors, increase the network lifetime and the continuous delivery and availability of services, collect complex data starting from diverse physical sensors, and decouple the physical sensors resources from the developed sensing application. The new expected types of cloud sensors for healthcare applications should also meet the requirements in terms of quality of service, security, and energy preservation.

The last issue is related to the need to cope with the increasing need of energy. In fact, the prominent need to make investigation in healthcare be proactive instead of postmortem, makes it necessary to develop novel techniques for the continuous collection and delivery of evidence, and the instantaneous reaction to potential digital crimes. These techniques will certainly increase the consumption of energy from the battery of medical equipment (wearable or mobile sensors), making it necessary to optimize the used algorithms to reduce their consumption of energy, or to find alternative sources of power. One of the potential used solutions is to use highly sophisticated batteries and power management systems, which will certainly increase the cost of the healthcare solution. Another, alternative and highly cheaper solution consists in using power harvesting or energy scavenging techniques to power the different tiny and autonomous sensors from many external sources such as solar power, thermal energy, and Radio Frequency energy. The ongoing and future research development should work on eliminating the battery from the energy-sensitive medical equipment and get rid of the weakest component in these devices. In some devices, such as pacemakers, the use of energy harvesting techniques would avoid doing surgical intervention by physicians to replace the patient's pacemaker each time its battery is depleted.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

In this chapter, we took interest in the postmortem investigation of digital crimes on healthcare applications. Firstly, we provided an overview of healthcare systems' architectures together with a detailed description of the involved components and their functions. We also reviewed and classified the most important types of healthcare applications proposed in the literature. Then, we identified and described a representative set of digital healthcare crimes. These crimes have been organized into four classes based on their goals and their impact on the patients' safety and privacy. We also discussed the impact of these crimes on the availability and the accuracy of the provided healthcare services. After identifying digital crimes on healthcare applications, we discussed the main requirements that should be fulfilled during the design of a postmortem investigation system tailored to these applications. A literature review of research works related to the use of theoretical techniques for the investigation of digital crimes on healthcare applications, and also the investigation of physical health crime scenes, was provided. Finally, we discussed a set of advanced issues that should be addressed during the design of an efficient postmortem investigation system capable of providing accurate results.

REFERENCES

Al-Qurishi, M., Al-Rakhami, M., Al-Qershi, F., Hassan, M. M., Alamri, A., Khan, H. U., & Xiang, Y. (2015). A Framework for Cloud-Based Healthcare Services to Monitor Noncommunicable Diseases Patient. *International Journal of Distributed Sensor Networks*, 2015.

Almadani, B., Bin-Yahya, M., & Shakshuki, E. M. (2015). E-AMBULANCE: Real-Time Integration Platform for Heterogeneous Medical Telemetry System. *Procedia Computer Science*, *63*, 400–407. doi:10.1016/j.procs.2015.08.359

Aquino-Santos, R., Martinez-Castro, D., Edwards-Block, A., & Murillo-Piedrahita, A. F. (2013). Wireless sensor networks for ambient assisted living. *Sensors (Basel, Switzerland)*, *13*(12), 16384–16405. doi:10.3390131216384 PMID:24351665

Astorga, J., Astorga, J. C., Jacob, E., Toledo, N., & Higuero, M. (2014). Securing access to next generation IP-enabled pacemakers and ICDs using Ladon. *Journal of Ambient Intelligence and Smart Environments, 6*(2), 157-177.

Carrier, B. D., & Spafford, E. H. (2006). Categories of digital investigation analysis techniques based on the computer history model. *Digital Investigation*, *3*, 121–130. doi:10.1016/j.diin.2006.06.011

Chen, H., Cheng, B. C., Liao, G. T., & Kuo, T. C. (2014). Hybrid classification engine for cardiac arrhythmia cloud service in elderly healthcare management. *Journal of Visual Languages and Computing*, *25*(6), 745–753. doi:10.1016/j.jvlc.2014.09.012

Cusack, B., & Kyaw, A. K. (2012). Forensic readiness for wireless medical systems. In *Proceedings of the 10th Australian Digital Forensics Conference* (pp. 21-32).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Dmitrienko, A., Hadzic, Z., Löhr, H., Sadeghi, A. R., & Winandy, M. (2013). Securing the access to electronic health records on mobile phones. *Biomedical Engineering Systems and Technologies*, *273*, 365–379.

Ellouze, N., Allouche, M., Ben Ahmed, H., Rekhis, S., & Boudriga, N. (2013). Securing implantable cardiac medical devices: use of radio frequency energy harvesting. In *Proceedings of the third international workshop on Trustworthy embedded devices* (pp. 35-42). 10.1145/2517300.2517307

Ellouze, N., Rekhis, S., Allouche, M., & Boudriga, N. (2014). Digital Investigation of Security Attacks on Cardiac Implantable Medical Devices. In *Proceedings of the Advanced Intrusion and Prevention Workshop in conjunction with the International Information Security and Privacy Conference* (pp. 15-30). 10.4204/EPTCS.165.2

Felisberto, F., Laza, R., Fdez-Riverola, F., & Pereira, A. (2015). A Distributed Multiagent System Architecture for Body Area Networks Applied to Healthcare Monitoring. *BioMed Research International, 2015*(2015).

Hei, X., Du, X., Wu, J., & Hu, F. (2010). Defending resource depletion attacks on implantable medical devices. In *Proceedings of the Global Telecommunications Conference* (pp. 1-5). 10.1109/GLOCOM.2010.5685228

Hu, F., Jiang, M., Celentano, L., & Xiao, Y. (2008). Robust medical ad hoc sensor networks (MASN) with wavelet-based ECG data mining. *Ad Hoc Networks*, *6*(7), 986–1012. doi:10.1016/j.adhoc.2007.09.002

Hui, J. W., & Culler, D. E. (2008). Extending IP to low-power, wireless personal area networks. *IEEE Internet Computing*, *12*(4), 37–45. doi:10.1109/MIC.2008.79

James, J. I., & Gladyshev, P. (2015). Automated inference of past action instances in digital investigations. *International Journal of Information Security*, *14*(3), 249–261. doi:10.100710207-014-0249-6

Kaur, P. D., & Chana, I. (2014). Cloud based intelligent system for delivering health care as a service. *Computer Methods and Programs in Biomedicine*, *113*(1), 346–359. doi:10.1016/j.cmpb.2013.09.013 PMID:24139021

Keppens, J., & Schafer, B. (2006). Knowledge based crime scenario modelling. *Expert Systems with Applications*, *30*(2), 203–222. doi:10.1016/j.eswa.2005.07.011

Ko, J., Lim, J. H., Chen, Y., Musvaloiu-E, R., Terzis, A., Masson, G. M., ... Dutton, R. P. (2010). MEDiSN: Medical emergency detection in sensor networks. *ACM Transactions on Embedded Computing Systems*, *10*(1), 1–29. doi:10.1145/1814539.1814550

Kumar, P., & Lee, H. J. (2011). Security issues in healthcare applications using wireless medical sensor networks: A survey. *Sensors (Basel, Switzerland)*, *12*(1), 55–91. doi:10.3390120100055 PMID:22368458

Li, C., Raghunathan, A., & Jha, N. K. (2011). Hijacking an insulin pump: Security attacks and defenses for a diabetes therapy system. In *Proceedings of the 13th IEEE International Conference on e-Health Networking Applications and Services* (pp. 150-156).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Magaña-Espinoza, P., Aquino-Santos, R., Cárdenas-Benítez, N., Aguilar-Velasco, J., Buenrostro-Segura, C., Edwards-Block, A., & Medina-Cass, A. (2014). WiSPH: A Wireless Sensor Network-Based Home Care Monitoring System. *Sensors (Basel, Switzerland)*, *14*(4), 7096–7119. doi:10.3390140407096 PMID:24759112

Maia, P., Batista, T., Cavalcante, E., Baffa, A., Delicato, F. C., Pires, P. F., & Zomaya, A. (2014). A Web platform for interconnecting body sensors and improving health care. *Procedia Computer Science, 40*, 135-142.

Mashima, D., & Ahamad, M. (2012). Enabling robust information accountability in e-healthcare systems. In *Proceedings of the third USENIX Workshop on Health Security and Privacy* (pp. 14-14).

Oh, S. E., Chun, J. Y., Jia, L., Garg, D., Gunter, C. A., & Datta, A. (2014). Privacy-preserving audit for broker-based health information exchange. In *Proceedings of the fourth ACM conference on Data and application security and privacy* (pp. 313-320). 10.1145/2557547.2557576

Rekhis, S., & Boudriga, N. (2011). Formal reconstruction of attack scenarios in mobile ad hoc and sensor networks. *EURASIP Journal on Wireless Communications and Networking*, (1): 1–17.

Savci, H. S., Sula, A., Wang, Z., Dogan, N. S., & Arvas, E. (2005, April). MICS transceivers: regulatory standards and applications [medical implant communications service]. In *Proceedings of the IEEE SoutheastCon* (pp. 179-182).

Srinivasan, V., Stankovic, J., & Whitehouse, K. (2008). Protecting your daily in-home activity information from a wireless snooping attack. In *Proceedings of the 10th international conference on Ubiquitous computing* (pp. 202-211). 10.1145/1409635.1409663

Steinfeld, R., Bull, L., Wang, H., & Pieprzyk, J. (2003). Universal designated-verifier signatures. Advances in Cryptology-Asiacrypt, 2894(2003), 523-542. doi:10.1007/978-3-540-40061-5_33

Tan, H., Hu, W., & Jha, S. (2011). A TPM-enabled remote attestation protocol (TRAP) in wireless sensor networks. In *Proceedings of the sixth ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks* (pp. 9-16). 10.1145/2069087.2069090

Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2013). Modeling crime scenarios in a Bayesian network. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law* (pp. 150-159). 10.1145/2514601.2514618

KEY TERMS AND DEFINITIONS

Attack Scenario: A series of malicious and legitimate actions executed on a digital system to harm its security. At some step in the scenario, a malicious action is executed to transform the system state from safe to unsafe (these states are already defined in a security policy).

Digital Crimes: Encompass all offences against the secrecy, integrity, and availability of electronic data and digital systems. When executed, they could lead to a set of serious threats ranging from economic to life threats.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Digital Forensic Investigation: The process of a postmortem analysis of a compromised system, to identify and retrieve digital evidence, understand the incident, identify attackers, and reconstruct potential attack. To conduct a forensic investigation, scientific methods can be used to prove/disapprove a claim about the incident in a court of law.

E-Health Systems: Healthcare systems which integrate the use of information and communications technologies to provide medical services to patients. Thanks to the integration of these technologies, a wide set of e-health applications, which improve the quality of life of high risk patients, were developed.

Implantable Medical Devices: A specific class of digital devices (e.g., pacemakers, defibrillators, insulin pumps), which are surgically implanted into a human's body, and used to treat chronic ailments (cardiac arrhythmia, diabetes, Parkinson).

Lethal Attacks: Security attacks that generally target medical devices and e-health systems. When executed, these attacks seriously harm the health status of the victim patient and lead to its death.

Physical Healthcare Crimes: Encompass all violations against the person's bodies. They generally cause serious physical damage (e.g., abuse, homicide) to the bodies of victim persons.

This research was previously published in Data Mining Trends and Applications in Criminal Science and Investigations edited by Omowunmi E. Isafiade and Antoine B. Bagula; pages 169-210; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 6
Critical Issues and Challenges

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 90
Bidirectional Spreading Activation Method for Finding Human Diseases Relatedness Using Well-Formed Disease Ontology

Said Fathalla
University of Bonn, Germany & Alexandria University, Egypt

Yaman M. Khalid Kannot
Arab Academy for Science, Egypt

ABSTRACT

The successful application of semantic web in medical informatics and the fast expanding of biomedical knowledge have prompted to the requirement for a standardized representation of knowledge and an efficient algorithm for querying this extensive information. Spreading activation algorithm is suitable to work on incomplete and large datasets. This article presents a method called SAOO (Spreading Activation over Ontology) which identifies the relatedness between two human diseases by applying spreading activation algorithm based on bidirectional search technique over large disease ontology. The proposed methodology is divided into two phases: Semantic matching and Disease relatedness detection. In Semantic Matching, semantically identify diseases in user's query in the ontology. In the Disease Relatedness Detection, URIs of the diseases are passed to the relatedness detector which returns the set of diseases that may connect them. The proposed method improves the non-semantic medical systems by considering semantic domain knowledge to infer diseases relatedness.

DOI: 10.4018/978-1-7998-1204-3.ch090

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

Usually, computers were used to perform tasks with none sort of intelligence or semantics thus there is a necessity to make computers as intelligent as human (Antoniou & Van Harmelen, 2004). The defect of non-Semantic applications is painted by the statement "lack of semantics" particularly when the talk is concerning information retrieval. The semantic web's goal is to extend automation in processing web-based data and to enhance the interoperability of web-based information systems (Payne, Terry & Lassila, 2004). The semantic web gives the meaning of knowledge making software systems can perform intelligent tasks instead of users (Berners-Lee, Hendler & Lassila, 2001) using a semantic knowledgebase called Ontology. Ontologies is a repository in which information are organized and used in semantic-based applications (Shadbolt, Berners-Lee & Hall, 2006). Below a group of definitions that may be used as a basis for mathematical formulation of the term ontology and its components (Fonseca, Davis & Câmara, 2003):

Definition 1: A term is a triple $\tau = [\eta, \delta, A]$, $\tau \in T$, where η is a string of characters containing the name of the term, δ is a string of characters containing its definition and A is a set of attribute domains A1, A2, ..., An, each associated to a value set Vi.

Definition 2: A relation $\varphi: T \rightarrow T$, $\varphi \in \Phi$: is a function from T to T such that for every term $\tau 1 \in T$, there is a term $\tau 1 = \varphi(\tau 1)$, $\tau 2 \in T$.

Definition 3: A semantic relation σ between two concepts is one of the relations among the set of semantic relations $\Sigma = \{$Hypernym, Hyponymy (is-a), Meronomy (part-of), Synonymy$\}$. In linguistics, a hyponym is a word whose semantics is enclosed inside that of another word which called its hyperonym or its hypernym. In other words, a hyponym has the same a type-of relationship with its hypernym. For instance, dog, cat and cow are all hyponyms of Mammals (their hypernym); that, in turn, could be a hyponym of animal (Fromkin, Rodman & Hyams, 2013).

Ontologies are often used in many systems as a semantic knowledgebase (Oberle, 2014). In the field of health informatics, ontology is used for the representing and organizing medical vocabularies. Collins and Loftus (1975) discussed in their research that the Spreading Activation (SA) runs on semantic networks (Anderson et al., 1983) and is used for information retrieval process (Burnset et al., 1981). It runs on a graph structure that comprises of a set of nodes connected by edges. The concepts are nodes which have an activation value and the relations between them are represented by edges between these nodes. An activation value is assigned to each node in the graph and then the algorithm spreads to the nodes with the higher activation value. The algorithm runs in a set of iterations and terminates when a stopping condition is reached.

The main objective of this work is to identify whether two human diseases are related to each other or not. If so, identify the set of diseases that connecting them. Finding a relatedness between diseases is important because the doctor can treat patient not only based on the symptoms they suffer from but treat the real cause of this disease which may be related to another disease that causes these symptoms therefore, the doctor can treat the cause disease not the symptoms.

This paper is organized as follows. Section 1 gives a brief overview of semantic web and ontologies. Section 2 reviews the literature. Section 3 describes the ontology for SAOO. Section 4 introduces the methodology of the proposed work. Section 5 illustrates the methodology by showing a running example. Section 6 describes the prototype used to implement the methodology. Finally, Section 7 presents conclusions and future work.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2. LITERATURE REVIEW

There is a considerable amount of literature on biomedical knowledge management and medical decision making. Biomedical knowledge available in the web is increasing considerably as most of the biomedical research papers are published online. Identifying named text features in user's query is a crucial procedure for efficient text analysis and matching. Semantic matching is a technique used to identify information which is semantically matched information in structured data. Many diverse solutions of match have been proposed so far. Some examples of individual approaches addressing the matching problem can be found in (Dhamankar et al., 2004) and (Do & Rahm, 2002). In their cutting-edge paper of 2004, Kanavos, Theodoridis and Tsakalidis (2014) proposed a metasearch engine over PubMed1. It provides accurate clustering of results according to their specific topic using a biomedical ontology, named MeSH[2] in addition to RxNorm[3] that is a US-specific terminology in medicine that semantically links clinical drugs vocabularies and support the interoperation between them. Ngo, Cao and Le (2010) proposed an ontology-based model for semantic annotation and searching using a vector space model. It takes advantage of ontological features of both named entities and WordNet vocabularies, and develops a spreading activation algorithm for query expansion. As anticipated, their experiments demonstrate that the proposed model is better than the solely keyword-based model and also better than the ones using only WordNet or named entities. In addition, it merges various ontologies to improve the semantic search process. De Maio et al. (2011) proposed a project named ODINO which uses a fuzzy knowledge approach for disease diagnosis that supports medical decision making. Medical knowledge is defined according with physician experts and represent relations between medical disease and its symptomatology. ODINO has three main features which are faceted search of diseases through taxonomy constraints, disease catalogue browsing and preliminary medical diagnosis.

Many efforts have been made to represent medical information as a semantic knowledgebase for further processing by semantic applications. Despite the existence of many medical ontologies, it is important reuse and integration of ontologies is to establish suitable mappings between their concepts. In recent years there has been growing interest in ontology matching and integration (Shvaiko and Euzenat, 2013; Euzenat, 2011). Shvaiko and Euzenat (2013) surveyed the state of the art of ontology matching and addressed some important challenges for ontology matching. In addition, they analyzed the results of recent ontology matching evaluations. Below some of the famous medical-related ontologies will be listed:

- Disease Ontology (DO)[4] is a biomedical resource of standardized common and rare disease concepts with stable identifiers organized by disease etiology;
- Vaccine Ontology (VO)[5] has been developed through collaboration with vaccine researchers and many national and international centers and programs. VO is represented in the Web Ontology Language (OWL) and edited using the Protégé-OWL. Currently VO contains more than 2000 terms and relationships (He et al., 2009);
- The Infectious Disease Ontology (IDO)[6] ontologies are designed as a set of interoperable ontologies that will together provide coverage of the infectious disease domain. At the core of the set is a general Infectious Disease Ontology (IDO-Core) of entities relevant to both biomedical and clinical aspects of most infectious diseases;
- The Ontology for Biomedical Investigations (OBI)[7] is an integrated ontology for the description of life-science and clinical investigations;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- The Ontology for General Medical Science (OGMS)[8] is based on the papers toward an ontological treatment of disease and diagnosis and on carcinomas and other pathological entities. OGMS was previously called the Clinical Phenotype Ontology.

3. THE ONTOLOGY FOR SAOO

The semantic web uses the resource description format data model (RDF) to describe data models, including concepts and relationships between them. The main purpose of developing ontology is to use it as a semantic knowledge base representation for identifying concepts in a specific domain and discover concepts semantics and the relationships between them (Fathalla, Hassan & Kholief, 2013). The authors of Köhler (2013) and Croft (2014) discuss in their research that the human disease data is a cornerstone of biomedical research therefore there's an enormous need for a consistent representation of human disease for robust data analysis (Colomb, 2007; LePendu et al., 2011). Creating a biomedical knowledgebase in the form of ontologies creates a rigid knowledgebase for semantic annotation of biomedical data through defined concepts and relations connecting them (Hoehndorf, Dumontier & Gkoutos, 2012).

The disease ontology (DO) used in this work has been developed as a standardized ontology for human disease. The objective of DO is to provide the biomedical community with convenient, reusable and robust knowledgebase of human disease concepts, phenotype characteristics and related medical vocabulary disease concepts. It is a collaborative effort of researchers at Northwestern University, Center for Genetic Medicine and the University of Maryland School of Medicine, Institute for Genome Sciences. DO does not include progression (early, late, metastasis, stages) or manifestations (transient, acute, chronic) of disease as part of the definition of disease (Kibbe et al., 2014). The human disease ontology is concentrate on representing predominant and rare disease concepts captured across biomedical resources with the mission of providing a disease interface between data resources through ongoing support of disease terminology needs. Major enhancements to the DO database since 2012 has been made including: the content of DO has had 192 revisions, including the addition of 760 terms and 32% of all terms now include definitions, improved data structure, new areas of community based efforts, new community partners and a switch to web ontology language (OWL). The current version of the DO website (version 1.0) provides a comprehensive resource to perform full-text searching on the DO as well as exploring and visualizing relationships between terms. The Disease Ontology database has been updated to the latest ontology as of 02-03-2017. The DO project has had a considerable influence on the development of biomedical resources, as evidenced by 291 Google Scholar citations to DO's paper in 2012 (Schriml et al., 2012). Table 1 shows the number of classes, properties, triples and the number of imported ontologies.

4. METHODOLOGY

The methodology of the proposed work is divided into two phases: Semantic matching and Disease relatedness detection as shown in Figure 1. In Semantic matching phase, semantically search for the user submitted diseases to identify whether these diseases are exist in the ontology using concepts that are semantically related to query concepts or not. The output of this phase is whether the diseases are found in the ontology or not. In case of found, the URIs of both diseases are retrieved. In the Disease related-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Do's statistics

Type	Local	Imported	Inferred	Total
Number of Triples	187957	89224	2390	279453
AnnotationProperty	25	0	0	25
Class	8720	25030	0	33750
DatatypeProperty	1	0	0	1
ObjectProperty	0	1	0	1
Ontology	1	5	0	6

ness detection phase, the URI for each disease is passed to the relatedness detector to find whether they are related or not. If they are related, the algorithm returns the set of diseases that connect them in the path from the first to the second.

4.1. Semantic Matching

The main objective of this phase is to identify diseases in the ontology which is semantically related diseases in the submitted query. It is assumed that, when searching for a concept, it is also important to get synonyms of that concept. For example, the synonyms of the disease "Hyperuricemia" are "Hyperuricaemia" and "Uricacidemia" therefore, diseases describing these concepts are also retrieved as well. The matching process is performed by in the following proposed algorithm:

```
Input: Ontology O, String disease1, String disease2
Output: URI of both diseases when matched, null if not.
α←false
β←false
For each (Resource R in O) do
α = Semantic_Matching (R, disease1)
If (α≠false) then
URI [0] = Resource.getURI(R);
Else
Print ("Disease 1 not found")
Return null
End if
β = semantic_matching (R, disease2)
If (β≠false) then
URI [1] = Resource.getURI(R);
Else
Print ("Disease 2 not found")
Return null
End if
End for
```

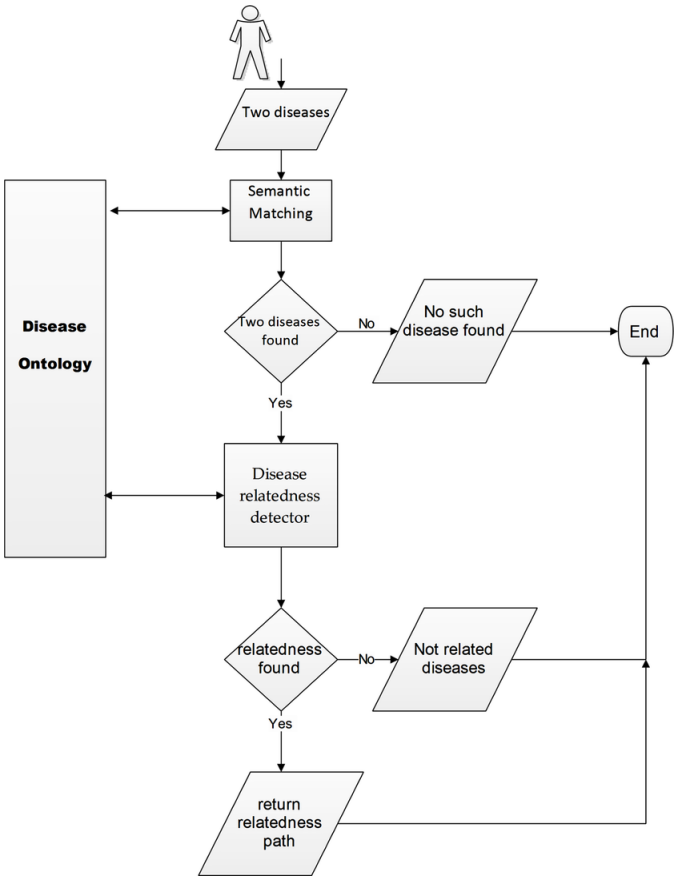

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The matched algorithm iterates through ontology resources and performs semantic matching between this disease and that resource. When matched, sets a flag α to true and then repeat this process for the second disease. Otherwise, the disease is not found in the ontology and returns null. Semantic matching conducts more analysis in matching process. This analysis may comprise term normalization; phrase analysis, word sense disambiguation, and the matching may be performed on phrase level and word sense level (Li & Xu, 2014).

4.2. Disease Relatedness Detection

In disease relatedness detection phase, the two diseases relatedness is detected by using bidirectional spreading activation on ontology graph. The idea of a bidirectional search is to cut back the search time by looking out forward from the beginning and backward from the goal at the same time. When the two search frontiers meet, the algorithm will reconstruct one path that extends from the beginning state through the frontier intersection to the goal. The merit of bidirectional search is its speed and it requires less memory (Poole & Mackworth, 2010).

Figure 1. Semantic matching process



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

5. RUNNING EXAMPLE

As a running example in this paper, a query includes two diseases ("Joubert syndrome with orofaciodigital defect" and "Meckel syndrome") will be used as a user submitted query. The results of relatedness detection may consist of a set of different diseases. As shown in Figure 2, the relatedness between "Joubert syndrome with orofaciodigital defect" and "Meckel syndrome" diseases are detected by the intersection nodes "Ciliopathy". The algorithm runs in the following steps:

1. Begin with two initial nodes "Joubert syndrome with orofaciodigital defect" and "Meckel syndrome" (marked in light gray);
2. The first cycle of activation process will active one more node from each of the starting node which is Joubert syndrome (from node 1) and Ciliopathy (from node 2);
3. The second cycle of activation process will active Ciliopathy and genetic disease diseases from the first and second path respectively;
4. The algorithm detects an intersection node Ciliopathy and stops.

The whole spreading process is summarized in Table 2.

6. EXPERIMENT

In this section, SAOO prototype is presented. This prototype exemplifies how spreading activation is used over ontology. SAOO automatically processes user's query in the following steps:

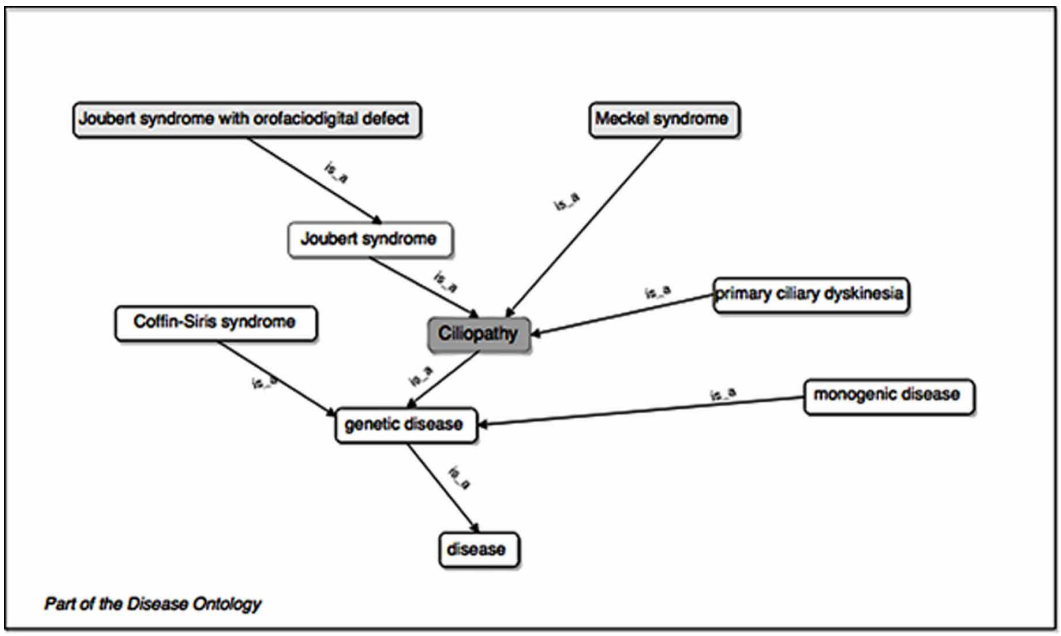
1. **User Input:** User submits the two diseases as a string;
2. **Query Preprocessing Including:** See the following:
   a. **Intelligent Tokenization and Stop List Elimination:** Tokenize the query string, by breaking it down into understandable segments. Usually a token is defined as an alpha-numeric string that separated by white space or punctuation. Then, eliminating all stop words. Those terms have little value in describing user's intention. A stop word list typically consists of those word classes that do not contain important significance to be used in the query, such as articles (an, a), conjunctions (and, thus), interjections (oh, but), prepositions (at, on), pronouns (he, she), and forms of the "to be" verb (is, are);
   b. **Stemming:** Stemming removes word suffixes: both inflectional suffixes (-s, -es, -ed) and derivational suffixes (*-able, -aciousness, -ability*) are stemmed;

Table 2. Summarization of the spreading process

1st Initial Node	Joubert syndrome with orofaciodigital defect
2nd Initial Node	Meckel syndrome
Path from node 1	Joubert syndrome- Ciliopathy
Path from node 2	Ciliopathy
Intersection	Ciliopathy

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The relatedness path between "joubert syndrome with orofaciodigital defect" and "meckel syndrome" diseases



3. **Formulate Semantic Query:** Take the preprocessed query string and formulate a semantic query using SPARQL and executes it against the ontology which performs the semantic matching process based on the inference rules (Fathalla, Hassan & El-Sayed, 2012). By incorporating these rules, semantically matched diseases could be found using Jena reasoners (Jena, 2015). These rules have been divided into three types: Transitive Rules, Symmetric Rules and Co-Reflexive Rules. In mathematical notation, these rules types defined in Table 3 (De Oliveira & Rodrigues, 2004). As an example of symmetric rule, if the disease "Hyperuricemia" is related to the disease "Hyperuricaemia" via hasSynonym property, then, it could be inferred that "Hyperuricaemia" must also be related to "Hyperuricemia" via hasSynonym property. In other words, if "Hyperuricemia" is a synonym of "Hyperuricaemia", then "Hyperuricaemia" must be a synonym of "Hyperuricemia". As an example of transitive rule, if a disease "Heart failure" has a synonym "Weak heart" and "Weak heart" itself has a synonym "CHF" then it could be inferred that "Heart failure" has synonym "CHF". Jena inference system is designed to permit a various inference engines or reasoners to be plugged into Jena. Such engines are able to infer logical consequences from a set of asserted facts or axioms together with any optional ontology information, axioms and rules associated with the reasoner. The main use of this technique is to support the use of languages which allow additional facts to be inferred from instance data and class descriptions such as RDFS and OWL. Available predefined reasoners included in the Jena distribution are (Jena, 2013):

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Inference rules types

Rule Type	Mathematical Notation
Transitive	$\forall a,b,c \in X: (aRb \wedge bRc) \Rightarrow aRc$
Symmetric	$\forall a,b \in X: aRb \Rightarrow bRa$
Co-reflexive	$\forall a,b \in X: aRb \Rightarrow b=a$

a. **Transitive Reasoner:** Provides support for storing and traversing property chains or transitivity axioms. This implements just the transitive and reflexive properties;
b. **RDFS Rule Reasoner:** Implements a configurable subset of the RDFS entailments;
c. **Generic Rule Reasoner:** This type of reasoners is very important because it supports user defined rules. Also, forward chaining, backward chaining and hybrid execution strategies are supported;

4. **Disease Relatedness Detection:** The diseases relatedness is detected by using bidirectional Spreading Activation on ontology graph if the diseases haven been found in the matching process.

7. CONCLUSION

This paper proposes a new method named SAOO (Spreading Activation over Ontology) in order to detect the relatedness between two human diseases by applying spreading activation algorithm based on bidirectional search technique over a well-formed disease ontology called DO. The main objective of this method is to identify whether two human diseases are related to each other or not. Moreover, identify related diseases that may connect them in a common path. The importance of the relatedness is that the doctor can treat patient not only based on the symptoms they suffer from but treat the real cause of this disease which may be related to another disease that causes these symptoms. Therefore, the doctor can treat the cause disease not the symptoms. The proposed method enhances the non-Semantic medical systems by considering semantic domain knowledge to infer diseases relatedness. For future work, authors are planning to find the relatedness between more than two diseases using multiple-goal search algorithms and provide the relatedness graph between them not only a simple path. In addition, a relatedness set is returned to the user which is a subset of the power set of the set of input diseases. In other words, which of these diseases is related to the others?

REFERENCES

Anderson, R. C., Liberta, A. E., Dickman, L. A., & Katz, A. J. (1983). Spatial variation in vesicular-arbuscular mycorrhiza spore density. *Bulletin of the Torrey Botanical Club*, *110*(4), 519–525. doi:10.2307/2996286

Antoniou, G., & Van Harmelen, F. (2004). *A semantic web primer*. MIT press.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*(5), 28–37. doi:10.1038cientificamerican0501-34 PMID:11341160

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Burns, E. C., Tanner, J. M., Preece, M. A., & Cameron, N. (1981). Final height and pubertal development in 55 children with idiopathic growth hormone deficiency, treated for between 2 and 15 years with human growth hormone. *European Journal of Pediatrics*, *137*(2), 155–164. PMID:7308226

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. doi:10.1037/0033-295X.82.6.407

Colomb, R. M. (2007). *Ontology and the Semantic Web* (Vol. 156). Frontiers in Artificial Intelligence and Applications.

Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., ... Jassal, B. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Research*, *42*(D1), D472–D477. doi:10.1093/nar/gkt1102 PMID:24243840

De Maio, C., Loia, V., Fenza, G., Gallo, M., Linciano, R., & Morrone, A. (2011, June). Fuzzy knowledge approach to automatic disease diagnosis. In Fuzzy Systems (FUZZ), 2011 IEEE International Conference on (pp. 2088-2095). IEEE. 10.1109/FUZZY.2011.6007498

De Oliveira, J. N. F., & Rodrigues, C. D. J. P. C. (2004, July). Transposing relations: from Maybe functions to hash tables. In *Proceedings of the International Conference on Mathematics of Program Construction* (pp. 334-356). Springer Berlin Heidelberg. 10.1007/978-3-540-27764-4_18

Dhamankar, R., Lee, Y., Doan, A., Halevy, A., & Domingos, P. (2004, June). iMAP: discovering complex semantic matches between database schemas. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data* (pp. 383-394). ACM. 10.1145/1007568.1007612

Do, H. H., & Rahm, E. (2002, August). COMA: a system for flexible combination of schema matching approaches. In *Proceedings of the 28th international conference on Very Large Data Bases* (pp. 610-621). 10.1016/B978-155860869-6/50060-3

Euzenat, J., Meilicke, C., Stuckenschmidt, H., Shvaiko, P., & Trojahn, C. (2011). Ontology alignment evaluation initiative: Six years of experience. *Journal on Data Semantics*, *XV*, 158–192.

Fathalla, S. M., Hassan, Y. F., & El-Sayed, M. (2012, October). A hybrid method for user query reformation and classification. In *Proceedings of the 2012 22nd International Conference on Computer Theory and Applications (ICCTA)* (pp. 132-138). IEEE. 10.1109/ICCTA.2012.6523559

Fathalla, S. M., Hassan, Y. F., & Kholief, M. (2013). Ontology-Based Approach for Automated Issue Classification in an Issue Tracking System. In *Proceedings of the 23rd International Conference on Computer Theory and Applications (ICCTA)*. IEEE.

Fonseca, F., Davis, C., & Câmara, G. (2003). Bridging ontologies and conceptual schemas in geographic information integration. *GeoInformatica*, *7*(4), 355–378. doi:10.1023/A:1025573406389

Fromkin, V., Rodman, R., & Hyams, N. (2013). *An introduction to language*. Cengage Learning.

He, Y., Cowell, L., Diehl, A. D., Mobley, H. L., Peters, B., Ruttenberg, A., ... Xiang, Z. (2009, July). VO: vaccine ontology. In *Proceedings of the 1st International Conference on Biomedical Ontology (ICBO 2009),* Buffalo, NY .

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hoehndorf, R., Dumontier, M., & Gkoutos, G. V. (2012). Evaluation of research in biomedical ontologies. *Briefings in Bioinformatics*. PMID:22962340

Jena, A. (2013). *Reasoners and rule engines: Jena inference support*. The Apache Software Foundation.

Kanavos, A., Theodoridis, E., & Tsakalidis, A. (2014, September). A pubmed meta search engine based on biomedical entity mining. In *Proceedings of the 2014 25th International Workshop on Database and Expert Systems Applications (DEXA)* (pp. 82-86). IEEE. 10.1109/DEXA.2014.32

Kibbe, W. A., Arze, C., Felix, V., Mitraka, E., Bolton, E., Fu, G., ... Parkinson, H. (2014). Disease Ontology 2015 update: An expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Research*. PMID:25348409

Köhler, S., Doelken, S. C., Mungall, C. J., Bauer, S., Firth, H. V., Bailleul-Forestier, I., ... FitzPatrick, D. R. (2013). The Human Phenotype Ontology project: Linking molecular biology and disease through phenotype data. *Nucleic Acids Research*. PMID:24217912

LePendu, P., Musen, M. A., & Shah, N. H. (2011). Enabling enrichment analysis with the human disease ontology. *Journal of Biomedical Informatics*, *44*, S31–S38. doi:10.1016/j.jbi.2011.04.007 PMID:21550421

Li, H., & Xu, J. (2014). Semantic matching in search. *Foundations and Trends in Information Retrieval*, *7*(5), 343-469.

Ngo, V. M., Cao, T. H., & Le, T. M. (2010, November). Combining named entities with wordnet and using query-oriented spreading activation for semantic text search. In Proceedings of the 2010 IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF) (pp. 1-6). IEEE. 10.1109/RIVF.2010.5633401

Oberle, D. (2014). How ontologies benefit enterprise applications. *Semantic Web*, *5*(6), 473–491.

Payne, T.R., & Lassila, O. (2004). Semantic web services. *IEEE Intelligent Systems*, *19*(1), 14–15.

Poole, D. L., & Mackworth, A. K. (2010). *Artificial Intelligence: foundations of computational agents*. Cambridge University Press. doi:10.1017/CBO9780511794797

Schriml, L. M., Arze, C., Nadendla, S., Chang, Y. W., Mazaitis, M., Felix, V., ... Kibbe, W. A. (2012). Disease Ontology: A backbone for disease semantic integration. *Nucleic Acids Research*, *40*(D1), D940–D946. doi:10.1093/nar/gkr972 PMID:22080554

Shadbolt, N., Berners-Lee, T., & Hall, W. (2006). The semantic web revisited. *IEEE Intelligent Systems*, *21*(3), 96–101. doi:10.1109/MIS.2006.62

Shvaiko, P., & Euzenat, J. (2013). Ontology matching: State of the art and future challenges. *IEEE Transactions on Knowledge and Data Engineering*, *25*(1), 158–176. doi:10.1109/TKDE.2011.253

Sourceforge. (n.d.). Jena, A Semantic Web Framework for Java. Retrieved from http://jena.sourceforge.net

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] Pubmed: http://www.ncbi.nlm.nih.gov/pubmed
[2] MeSH: https://www.ncbi.nlm.nih.gov/mesh
[3] RxNorm: https://www.nlm.nih.gov/research/umls/rxnorm/
[4] Disease Ontology: http://www.disease-ontology.org
[5] Vaccine Ontology (VO): http://www.violinet.org/vaccineontology
[6] Infectious Disease Ontology: http://infectiousdiseaseontology.org/page/Main_Page
[7] Ontology for Biomedical Investigations: http://obi-ontology.org/page/Main_Page
[8] Ontology for General Medical Science: https://bioportal.bioontology.org/ontologies/OGMS

This research was previously published in the International Journal of Computers in Clinical Practice (IJCCP), 2(1); edited by Athina A. Lazakidou; pages 45-58; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 91
An Investigation Into the Efficacy of Deep Learning Tools for Big Data Analysis in Health Care

Rojalina Priyadarshini
KIIT University, India

Rabindra K. Barik
KIIT University, India

Chhabi Panigrahi
University of Rajastan, India

Harishchandra Dubey
The University of Texas at Dallas, USA

Brojo Kishore Mishra
C. V. Raman College of Engineering, India

ABSTRACT

This article describes how machine learning (ML) algorithms are very useful for analysis of data and finding some meaningful information out of them, which could be used in various other applications. In the last few years, an explosive growth has been seen in the dimension and structure of data. There are several difficulties faced by conventional ML algorithms while dealing with such highly voluminous and unstructured big data. The modern ML tools are designed and used to deal with all sorts of complexities of data. Deep learning (DL) is one of the modern ML tools which are commonly used to find the hidden structure and cohesion among these large data sets by giving proper training in parallel platforms with intelligent optimization techniques to further analyze and interpret the data for future prediction and classification. This article focuses on the use of DL tools and software which are used in past couple of years in various areas and especially in the area of healthcare applications.

DOI: 10.4018/978-1-7998-1204-3.ch091

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Machine learning (ML) is a field of study which learns the behaviour of functional relationship among data without considering the detail underlying characteristics present in them. These techniques are applied to derive the functional attributes of underlying data which do not contain precise definition about them. In real cases, the datasets which are available do not have a precise definition and functional co-relation. So, ML algorithms are used appropriately for these types of data which are capable to infer the useful features of the data and their functional co-relationship so that they can be used in the problems of classification, clustering and prediction etc. Deep learning (DL) is a field of study under the umbrella of Machine learning (ML) which is constituted on the algorithms inspired by the structure and utility of human brain. This study is coherent to artificial neural network (ANN) and can be called as advanced ANN having higher set of problem solving capability and utility. DL refers to a special type of ANN which is constituted by multiple numbers of hidden layers. The hidden layers are responsible for handling nonlinearity and complexity of underlying data. These are mainly useful for giving training to network exhibiting complex co relational functions and behaviour and are mostly suitable for large data sets. Due to this reason these are very successfully applied to big data domain. In this contemporary era, data is highly scalable and growing exponentially. There is a parallel growth in data reproduction, availability and access and abundance high speed computers. So, the need of the hour is to develop high speed large neural network which are scalable enough to deal with plethora amount of big data. The input to the conventional shallow learning algorithms are known features from the data, but DL algorithms possess the potential to extract the representative feature available from the raw data whose sources are social media, transactional data, business houses, etc.

In most application of data like classification, clustering and prediction the common work flows for machine algorithms are built on some basic steps like: (i) data cleaning (ii) pre-processing (iii) training which build the learning model and (iv)testing which will evaluate the built model. For these problems the input space contains a set of input data sample x_i ($i = 1, 2, ..., n$; where n is the number of input instances which is a vector of 'm' dimension where m is the number of features) and output y_i . In supervised algorithms the artificial model learns to derive a mapping function $f(x) = y$ from a priory known pair of (x_i, y_i). After the model is duly trained on the basis of known input-output pairs (x_i, y_i), it is evaluated against the unknown samples to infer the target output which is called as testing.

In the case of unsupervised model, the model has to explore the pattern exhibiting in input x_i to find out y_i. In the training phase output y_i is not provided to the model, it has to find itself. Clustering and principal component analysis are some of the examples of unsupervised learning.

Deriving feature sets x_i to build an input dataset is very time consuming and problem specific. These are less available and difficult to create for every problem. This is the reason for unsuitability of supervised and unsupervised algorithm in all problems. But deducting the principal feature from a data set is directly associated with the performance of a learning algorithm. If features are correctly fetched from the data, then learning will take less time and the model shows higher level of accuracy. In the other hand for high dimensional data containing big data components, it is almost impossible to find out all the informative and relevant features from the raw data. This is a major limitation behind using conventional machine learning algorithms in big data possessing high dimensional feature data. This limitation is overcome by the emergence of deep neural network (Bengio et al., 2013; LeCun et al., 2015). Deep neural network contains many layers containing more than one hidden layer (Oliveira et al., 2016). The input layer takes the high dimensional raw data and extracts the relevant representive

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

features from them. These abstract features are regenerated in subsequent hidden layers to incorporate the complex and nonlinear functions exhibited in high dimensional big data. The potential of deep learning algorithms to handle high dimensional raw data make them popular in the field of data analysis to deal with data having components like volume, variety, veracity, versatility of data (Aye & Thein, 2015).

BACKGROUND

Apart from health care, the application domains of DL are in various sphere such as: speech and audio processing, natural language processing, acoustic modelling, speech synthesis, language modelling, object recognition, computer vision, etc. (LeCun et al., 2015). Some specific task may include colorization of black and white images, classification of objects in still photographs, auto generation of handwriting, automatic system for gaming, generation of appropriate captions for images etc. In last couple of years DL has been widely used in the area of public health care and life science (NVIDIA corp. (2016). The data analyst from NVIDIA has recently disclosed in a report about the estimation of possession of AI components is 80% in applications by the year 2020. They have discussed about the recent advancement of medicine world and pointed the role of DL in bringing transformation in medicine world. Some of the focused areas in health care which indulge DL are discussed below.

Mining Clinical Data for Faster Treatment

The medical records of a patient which may include prescription details, medical images, MRI scans and clinical test reports are the useful information for a doctor. Maintain all these data for all the patient is a cumbersome job. Extracting useful information on time will help the doctor for diagnosis and detection of a disease. In these cases, GPU accelerated DL tools can be used to process these data and finding relevant information of a patient data for quicker treatment.

Quicker Diagnosis Process

Doctors rely greatly on medical images like MRIs, CT Scans, X-Ray reports, Ultra sound reports, ECG signals to detect and diagnose some diseases. Sometime these data are hard to analyse and consume a lot of time. In this case DL can be used as an automatic tool to analyze the data, which can further help the medical practitioners to make the diagnosis process faster.

Prescribing Personalized Medicines

The genomic data are responsible for causing disease in living body. Identifying and studying the genomic data causing disease is very difficult. Because genomic data are voluminous and extracting genetic parameters from huge genomic data are computation intensive jobs and require faster processing. DL can be employed to handle voluminous genetic data to do processing on these to mine information of disease causing genetic materials. So that personal medicines can be prescribed to individual patient according to the genetic structure and analysis. A GPU accelerated project taken over by a class of researchers from University of Toronto are working on a project named as" genetic interpretation engine". Their work is focused on identifying cancer causing genes present in individual patient.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Helping Disabled People

Visually impaired people face a lot of difficulties while doing routine tasks such as road crossing, recognising people, reading tags and labels of different products. For this segment of people wearable devices are being developed containing embedded prior programmed chips. DL algorithms with the help of other techniques like image processing, computer vision is employed as embedded programs to perceive and understand the surroundings and transform these perceptions to words so that the blind people can get help to know their surrounding and accomplish the task.

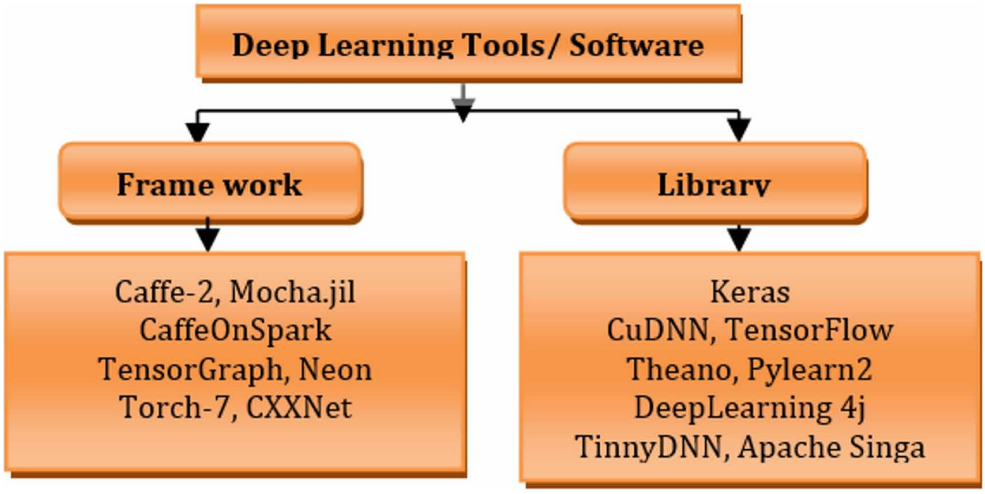
Prediction of Pharmacological Properties of Drugs

Experimental gene expression data collected from human cells are applied as input data for studying the functional changes by using chemical and drugs. These study focuses on the reaction of different newly designed drugs by exploiting them to gene expression data. These types of data are voluminous and hard to compute with conventional methods. Deep neural network is used in these cases to forecast properties of drugs on various diseases (Mamoshina et al., 2016).

OBJECTIVE OF PRESENT WORK

Different types of DL architectures are: Deep neural network; Deep Auto-encoder; Deep Belief network; Deep Boltzmann machine; recurrent neural network; convolution neural network. Among these the widely used is deep neural network. The available software to implement and deploy DL models do not support all mentioned types. Some tools are specific, and some are generic. Some have fixed architecture, and some have customizable modules. The deployment of DL models can be done through two parts: (i) Frame work or (ii) Library. These are responsible for model different layers as a building block of the network and the architecture of the network is fully dependent on the problem domain. Figure 1 presents the available DL libraries and framework.

Figure 1. List of DL framework and libraries



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

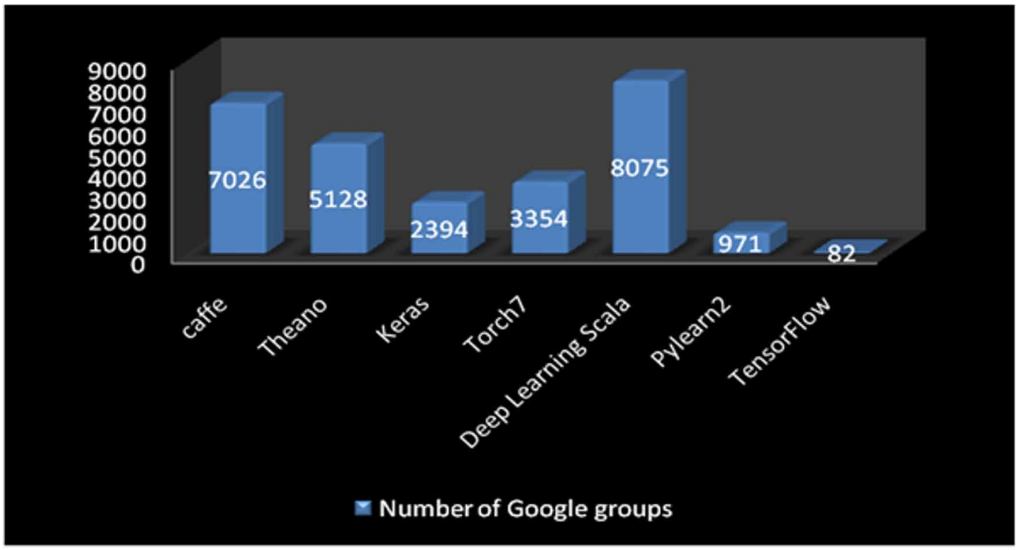
The present work focuses on analysis of the deep learning tools being used and can be applied in the field of health care system. The parameters like their make, platform, execution environment, interface support, application domain, implementation details, distributiveness, parallel environment support are the main points which are investigated in this paper. The limitations of these tools are studied and depicted. The results are taken from studied literatures from reputed article, industry white papers and product documentations.

Some of widely accepted tools found in literature survey are depicted below. The statistics of some of these tools' popularity is quantized by number of Google groups using these tools are shown in Figure 2.

Caffe

This is a deep learning framework which is developed by Berkly AI Research which is meant to facilitate a complete package for training, testing, adjusting the deep learning model parameters (CenterBerkeley, 2016). It supports modularity, so extension over new version would be easy. It has a clear isolation between representation and execution layer. So, switching between CPU and GPU is very easy and it just needs to call a function. Python and Matlab languages can be employed to create networks and for data classification, thereby supporting interface to bind these. It provides pre-trained reference models which help new researchers to quick start with it. In caffe through configuration file, network architectural parameters are set, and the network can be trained and tested through command line. This helps users not to write detail coding to configure and deploy the model. This tool is widely used for image recognition and computer vision task. It is applicable in medical image analysis. Some limitations of caffe include, it is not well designed for recurrent network; customization can be made with C++ only (Jia et al., 2014).

Figure 2. The number of google groups working with DL software (as on 14-05-2017)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Theano

Theano is built on a set of python library, is a declarative programming model aiming to provide research support in deep learning. Its native code is C++ and Python and innovated by University of Montreal. The core part of machine learning application's execution is doing complex matrix operations. It provides a rich library to solve this mathematical expression with the help an intelligent compiler which uses heuristics to optimize the expressions and swiftly handle multidimensional arrays. It also has the support of GPU devices which run the expressions faster and with a higher degree of accuracy. It can be customized according to user application need. One drawback of Thiano is that, the compilation time is more while dealing with larger datasets. Because the network model is produced as a computational graph and reconverted to native code through the optimized compiler (Universite de Montreal, 2016).

Pylearn2

It is a library constituting of machine learning libraries built up in LISA at University de Montreal. Since this is created for researchers, it is having its own target users. The users must have expertise on machine learning algorithms to understand the code walk through. It can execute symbolic expressions directly. It is built on top of Theano. This library was aiming to be utilized for academic and scientific use. But it is a customizable module thus can be applied for any real data mining task like classification, prediction etc. The only limitation is that customization modules must be written in C++ (Goodfellow et al., 2013).

CaffeOn Spark

It integrates the features of DL framework Caffe and big data framework apache spark and apache hadoop. It is an open source framework that can create a distributed DL network with a cluster of CPU and GPU servers [Andy Feng, Jun Shi and Mridul Jain]. Unlike some other DL framework need an isolated cluster for deep learning, which requires large datasets to be migrated among cluster nodes which is leading to increase system latency and latency, here the cluster is built not separately and is embedded in Spark application program, by which unwanted data migration among cluster nodes can be prevented (Kim et al., 2016).

NVIDIA cuDNN[]

cuDNN a collection of primitives to implement standard routines of deep neural network to create, train and evaluate the network. It is a GPU accelerated framework which provides support to create software applications build on created neural network models. It contains number of tensor transformation functions which are helpful for repetitive computation done in ML algorithms. It can be integrated with caffe and Baidu (Chetlur, 2014).

keras

It is a Python library to implement deep learning which makes use of either Thiano or TensorFlow as backend. It runs in python 2.7 or 3.5 versions. This can be used to design deep learning models both for research and development. It is an open source tool. The main guiding principle on which it is built

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

is modularity and extensibility. Its native is Python and no other language is supported. It is used as a binary and multiclass classifier. It can also be used for regression analysis (Franois Chollet, 2016).

Deep Learning 4j

Deep learning 4j java based open source library for creating and deploying deep neural network. It works in the company of Spark and Hadoop. Its native is Java having Apache 2.0 license, so all JVM languages such as Scala or Closurie are compatible with it. This is written in C++. It can be integrated with Haddop and Spark and can execute on various other backends. It gives a plug-and-play solution instead of building configuration. A part of Deep learning 4j, ND4j is a vectorization tool containing a rich matrix library to do core machine learning tasks. This tool can be used to create recurrent network and supports reinforcement learning. Unlike other tools it is solely built for commercial use in business house. This framework is distributed and supports multi-threaded programming model, thus is capable to run in parallel environment. It is a complete package to provide all sorts of application to model a deep neural network possessing smart n-dimensional array class (Skymind, 2016).

Mocha.jil

It is a DL library developed at MIT for a new programming language Julia, especially designed to do scientific and numerical computation. Its native code is Julia and can interface with Julia only. It can work with Julia backend, but it supports GPU backend and can be integrated with Nvidia libraries like cuDNN. It is highly modular and can be easily extended. It does not require configuration file for describing network architecture. Network architecture can be directly written by Julia code. It contains some pre-trained models to do a quick start. The only difficulty with this is it is compatible with Julia only (Ravi et al., 2017).

Torch7

Torch7 (MIT, 2015) was first built up in the University of New York. Its foundation is on a special scripting language called LuaJIT. Lua is chosen because it can easily be integrated with C language. By placing appropriately, the existing modules the deep Networks could be formulated. As based on a scripting language the model is not required to be compiled. This makes it suitable for quick prototyping if compared to Theano (Collobert, 2011). This one is used in facebook AI lab and Google deepmind. A possible downside of this tool is that the user has to be familiar with the LuaJIT scripting language. Also, LuaJIT is not good at creating custom recurrent networks.

TensorFlow

Like Theano, TensorFlow is also an open source library. It is designed to do numerical computation by using data flow graphs as in case of neural network. It was innovated by the researchers of the Google Brain Team at Google's Machine Intelligence laboratory. Unlike Thiano, TensorFlow is based on distributed computing, so well fit to work with many GPUs and CPUs in clusters. It has an inherent support for heterogeneous and parallel computation (Google, 2016). This tool is developed in C++ and has

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

python compatibility. One demerit of TensorFlow is that, pre-trained model is not available in a large proportion. It is used in handling biological data in bioinformatics to solve different problems like gene selection, prediction etc. (Rampasek, 2016).

TensorGraph

It is an easier and simpler framework for creating any graphical model based on TensorFlow. It contains three types of nodes: StartNode which takes input for the graph, HiddenNode for creating the hidden layers and EndNode used for generating the outputs of the model (Xu, 2013).

Tinny-dnn

It is a header file written in C++ containing some deep learning library. It can be easily integrated with any compatible application to build network. The network can be trained and tested by using suitable methods. It contains several activation functions such as tanh, sigmoid, softmax, leaky relu etc. (Angelova, 2015).

MXNET

MXNET is rapid, and compact distributed deep learning framework basing on MShadow. MShadow is a lightweight CPU/GPU Matrix Template Library in C++/CUDA. The goal of mshadow is to provide device invariant and simple matrix library for ML work. It Supports CPU/GPU/Multi-GPU and distributed system. It is a lightweight and highly modular C++/CUDA neural network toolkit, which can interface with Python and Matlab for training and prediction (Chen et al., 2015).

Neon

This is an open source python based framework developed at Intel, used to create DL models. It is used to develop convolution network with multi-GPU support. It does optimization at assembler level and data loading level. It has rich visualizations library to well understand DL process. It involves object oriented approach for implementing DL constructs such as layers, learning rules, activation functions etc. It can be used multi-layer perceptrons, recurrent neural networks and auto encoders. It is included with pre-trained examples for image captioning and sentiment analysis work (Nervana Systems, 2016).

RESULT AND DISCUSSION

The analysis and investigations done throughout the above studies are presented in Table 1, Table 2 and Table 3. Table 1 gives a comparative analysis of DL software and their characteristics details. Table 2 gives the common areas in health care system which uses DL software. It also focuses some specific application areas for which solution domain uses machine learning and deep learning algorithms along with the input data used for the problem. The prime areas are health informatics (Simonyan et al., 2014) Bioinformatics (Ravi et al., 2017; Chrimes, et al., 2017), public healthcare and Medical imaging. Table 3 gives brief comparison about the tools along with some inferred observations regarding their pros and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

cons and the areas of application. It can be observed from the literature that, the use of DL in health care domain and other is increased exponentially in recent past. Most of the popular tools use Python as the interfacing language. The commercial available tools are less. One of the examples is Deep learning 4j. Most of the tools are open source and can be customizable for research and scientific applications. With the increase of structure and volume of data, the tools are also evolving. There are certain limitations existing for using DL in health care can be noted here (Chen et al., 2014):

- The field of DL is new to researchers and scientist. There are many questions still lying with DL to be answered.
- For training deep, neural network a large amount of training data is required. This is causing a major problem to use DL with the present health care system. Because the electronic data keeping system is yet to be achieved in this sector.
- The visualization of DL technique is still remains a black box for many researchers.

Table 1. Comparison of DL software

Software	Developed & Maintained By	Open source	Native Code	Support Interface	Nvidia cuDNN Support	Presence of preexisting training model
Caffe	Berkley AI Research Lab	Yes	C++	MATLAB, Python	Yes	Yes
Theano	University of Montreal	Yes	Python	Python	Yes	Yes
Pylearn2	LISA at University de Montreal	Yes	Python	Python	Yes	Yes
CaffeOn Spark	Yahoo Research Lab	No	Not available (NA)	Apache Hadoop, Spark, Scala	Yes	NA
cuDNN	Nvidia	Yes	C	Can be integrated with other existing DL models	Yes	NA
MXNet	DMLC Group	Yes	C++	R, Python, Julia, and C++	Yes	Yes
Keras	François Chollet	Yes	Python	Python	Yes	Yes
Mocha-jil	MIT	Yes	Julia	Julia	Yes	Yes, but very few
Deep learning 4j	SkyMind	Yes	Java	Java	Yes	Yes
Torch7	Ronan Collobert et al.	Yes	Lua	Lua, LuaJIT, C	Yes	
TensorFlow	Google	Yes	C++	Python	Yes	No
TensorGraph	Google	Yes	NA	Python	Yes	No
TinnyDNN	Tiaga Nomi	Yes	C++	C++ supporting languages	NA	No
Neon	Intel Nervana	Yes	Python	Python	Yes	Yes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Summary of areas and application in health care with input

Area	Input Data	Applications
Clinical Informatics	Digital clinical data, Medical records	Disease prediction Human health monitoring Early Notification on abnormality
Public Healthcare	Geospatial images Social media personal data captured from smart phone Text data	Air Pollution detection Lifestyle disease controller Demographic information prediction Detection of Infectious diseases epidemics
Medical Imaging	MRI/CITI Scan reports X-Ray images Endoscopy pictures	Neural cells classification Healthy/unhealthy tissue classification Tumor detection Blood hemorrhage detection Cell clustering
Bioinformatics	DNA Samples Gene sequence Microarray Data	Drug designing Cancer diagnosis Gene classification Gene selection

Table 3. DL software with their pros and cons

Software	Programming Model	Suitable for	Pros & Cons
Caffe	Declarative	Image recognition Computer vision Speech recognition	**Pros** User does not have to write detail code. Pre-trained models are present to do a quick start **Cons** Cannot be used to create recurrent network Customization can be made only through C++.
Theano	Declarative	Numerical computation	**Pros** Good for tensor based mathematical expressions Compiler can optimize expressions **Cons** Compilation time is more for larger datasets
Pylearn2	Declarative	Classification Prediction	**Pros** Quicker computation **Cons** Customization can be made only through C++. User should have good knowledge on DL to understand the code execution.
Keras	Emparative	Prediction Image classification Large scale image recognition Feature extraction	**Pros** Easily extendible **Cons** Only interface is Python
Deep learning 4j	Object oriented programming	Object recognition in Social networks for Photo tagging	**Pros** Can import neural net models from Tensorflow, caffe, Torch and Theano. It is distributed Complete package suitable for commercial use
Torch7	Emparative	Used in Facebook AI research IBM Research lab	**Pros** Easily integrated with C No Compilation is needed **Cons** User has to be familiar with LuaJIT
TensorFlow	Emparative	Image classification Drug Designing Information Retrieval	**Pros** Distributed and inherently paralle **Cons** No pre-trained libraries
Neon	Object oriented	Image captioning Sentiment analysis	**Pros** Optimization at assembler and data loading level Rich graphical & visualization properties

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Nvidia cited that use of ML algorithm in health will change the dimension of current health care system. Earlier data were limited in size. So, for mining information from data and using the same results to let a machine to estimate and predict something can be established by using simple neural networks having different learning rules. With the explosive increase of data, its dimension and structural composition are changing. The complexity and nonlinearity has been changed.

So, present models are not able to compute these data. DL came into picture few years ago. Their architecture of having more intermediate computing layers is designed to handle the complexities and nonlinearity of big data. Some widely used tools like Caffe, Thiano, Deep learning 4j, Tensor Flow are discussed in the study by taking into consideration, some of their technical specifications, their strength and weakness (Min et al., 2016; Deng & Yu, 2014; Poonia et al., 2016; Schmidhuber et al., 2015). The health-related application areas are discussed along with the specific problem, which are using DL tools and techniques. Most of the tools are designed in the view of scientific and academic use. Full-fledged commercial tools are yet to come. Maximum of the available tools are open source and can be customizable according to requirement. As the technique is its early stage, some concrete hypothetical theories are yet to be evolved.

REFERENCES

Angelova, A., Krizhevsky, A., Vanhoucke, V., Ogale, A. S., & Ferguson, D. (2015, September). Real-Time Pedestrian Detection with Deep Network Cascades. In BMVC (pp. 32-1). doi:10.5244/C.29.32

Aye, K. N., & Thein, T. (2015). A platform for big data analytics on distributed scale-out storage system. *International Journal of Big Data Intelligence*, *2*(2), 127–141. doi:10.1504/IJBDI.2015.069088

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. doi:10.1109/TPAMI.2013.50 PMID:23787338

CenterBerkeley. “Caffe,”2016.[Online].Available:http://caffe.berkeley vision.org/

Chen, T., Li, M., Li, Y., Lin, M., Wang, N., Wang, M., . . . Zhang, Z. (2015). Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems. arXiv:1512.01274.

Chen, X. W., & Lin, X. (2014). Big data deep learning: Challenges and perspectives. *IEEE Access: Practical Innovations, Open Solutions*, *2*, 514–525. doi:10.1109/ACCESS.2014.2325029

Chetlur, S., Woolley, C., Vandermersch, P., Cohen, J., Tran, J., Catanzaro, B., & Shelhamer, E. (2014). cudnn: Efficient primitives for deep learning. arXiv:1410.0759

Chrimes, D., Kuo, M. H., Moa, B., & Hu, W. (2017). Towards a real-time big data analytics platform for health applications. *International Journal of Big Data Intelligence*, *4*(2), 61–80. doi:10.1504/IJBDI.2017.083131

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, *12*(Aug), 2493–2537.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and Trends in Signal Processing*, *7*(3–4), 197-387.

Franois Chollet. (2016). Keras. Retrieved from https://keras.io/

Goodfellow, I. J., Warde-Farley, D., Lamblin, P., Dumoulin, V., Mirza, M., Pascanu, R., . . . Bengio, Y. (2013). Pylearn2: a machine learning research library. arXiv:1308.4214

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... Darrell, T. (2014, November). Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia* (pp. 675-678). ACM. 10.1145/2647868.2654889

Kim, H., Park, J., Jang, J., & Yoon, S. (2016). DeepSpark: A Spark-Based Distributed Deep Learning Framework for Commodity Clusters. arXiv:1602.08191

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi:10.1038/nature14539 PMID:26017442

Mamoshina, P., Vieira, A., Putin, E., & Zhavoronkov, A. (2016). Applications of deep learning in biomedicine. *Molecular Pharmaceutics*, *13*(5), 1445–1454. doi:10.1021/acs.molpharmaceut.5b00982 PMID:27007977

Min, S., Lee, B., & Yoon, S. (2016). Deep learning in bioinformatics. *Briefings in Bioinformatics*. doi:10.1093/bib/bbw068 PMID:27473064

NervanaSystems. (2016). Neon. Retrieved from https://github.com/NervanaSystems/neon

NVIDIA corp. (2016). Nvidia dgx-1. Retrieved from http://www.nvidia.com/object/deep-learning-system.html

Oliveira, T. P., Barbar, J. S., & Soares, A. S. (2016). Computer network traffic prediction: A comparison between traditional and deep learning neural networks. *International Journal of Big Data Intelligence*, *3*(1), 28–37. doi:10.1504/IJBDI.2016.073903

Poonia, P., Jain, V.K. & Kumar, A. (2016). Deep Learning: Review.

Rampasek, L., & Goldenberg, A. (2016). Tensorflow: Biology's gateway to deep learning? *Cell Systems*, *2*(1), 12–14. doi:10.1016/j.cels.2016.01.009 PMID:27136685

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, *21*(1), 4–21. doi:10.1109/JBHI.2016.2636665 PMID:28055930

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, *21*(1), 4–21. doi:10.1109/JBHI.2016.2636665 PMID:28055930

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. doi:10.1016/j.neunet.2014.09.003 PMID:25462637

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556

Skymind. (2016). Deeplearning4j. Retrieved from http:// deeplearning4j.org/

Tensorflow. (n.d.). Retrieved from https://www. tensorflow.org/

Universite de Montreal. (2016). Theano. Retrieved from http://deeplearning.net/software/theano/

Xu, S. (2013). Semi-supervised Tensor Graph-optimized Linear Discriminant Analysis for Two-dimensional Face Recognition. *JSW*, *8*(8), 2047–2051. doi:10.4304/jsw.8.8.2047-2051

Zhang, C., Li, P., Sun, G., Guan, Y., Xiao, B., & Cong, J. (2015, February). Optimizing fpga-based accelerator design for deep convolutional neural networks. In *Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays* (pp. 161-170). ACM. 10.1145/2684746.2689060

This research was previously published in the International Journal of Grid and High Performance Computing (IJGHPC), 10(3); edited by Emmanuel Udoh, Ching-Hsien Hsu, and Mohammad Khan; pages 1-13; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 92
Real Time Analysis Based on Intelligent Applications of Big Data and IoT in Smart Health Care Systems

Mamata Rath
C.V. Raman College of Engineering, India

ABSTRACT

Currently, there is an expanding interest for additional medical data from patients about their healthcare choices and related decisions, and they further need investment in their basic health issues. Big data provides patients presumptuous data to help them settle on the best choice and align with their medicinal treatment plan. One of the very advanced concepts related to the synthesis of big data sets to reveal the hidden pattern in them is big data analytics. It involves demanding techniques to mine and extract relevant data that includes the actions of piercing a database, effectively mine the data, query and inspect the data and is committed to enhance the technical execution of various task segments. The capacity to synthesize a lot of data can enable an association to manage data that can influence the business. In this way, the primary goal of big data analytics is to help business relationships to have enhanced comprehension of data, and subsequently, settle on proficient and very much educated decisions. Big data analytics empowers data diggers and researchers to examine an extensive volume of data that may not be outfit utilizing customary apparatuses. Big data analytics require advances and statistical instruments that can change a lot of organized, unstructured, and semi-organized data into more reasonable data and metadata designed for explanatory procedures. There is tremendous positive potential concerning the application of big data in human health care services and many related major applications are still in their developmental stages. The deployment of big data in health service demonstrates enhancing health care results and controlling the expenses of common people due to treatment, as proven by some developing use cases. Keeping in view such powerful processing capacity of big data analytics in various technical fields of modern civilization related to health care, the current research article presents a comprehensive study and investigation on big data analytics and its application in multiple sectors of society with significance in health care applications.

DOI: 10.4018/978-1-7998-1204-3.ch092

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

The use of big data in data analysis and synthesis have been utilized as a part of explanatory instruments must find examples, patterns, and relationships over an assortment of time skylines. In the wake of breaking down the data, these instruments envision the discoveries in tables, diagrams, and spatial outlines for proficient decision making. Along these lines, big data investigation is a genuine test for some applications due to data unpredictability and the adaptability of basic calculations that help such procedures acquiring supportive data from big data investigation is a basic issue that requires adaptable logical calculations and strategies to return all around planned outcomes, though current methods and calculations are wasteful to deal with big data analytics (Rath et al., 2019). In this manner, huge framework and extra applications are important to help data parallelism. In addition, data sources, for example, rapid data stream got from various data sources, have diverse configurations, which makes coordinating different hotspots for analytics arrangements basic. Subsequently, the test is centered around the execution of current calculations utilized as a part of big data investigation, which isn't rising directly with the fast increment in computational assets.

In general, when the term "big data," is coined people instantly consider large data volumes. On the off chance that these were the main purpose behind health care services to receive the better approach for putting away data, they could oversee without, in light of the fact that the majority of them could contain what they have in a strong social database (Sun et al., 2012). In 2001, Doug Laney portrayed the "3 Vs" of large and complex data as "Volume, Velocity and Variety." While medicinal services could profit by every one of the three, the accentuation ought to be on assortment. This is a pattern experienced by human services, as well as by all enterprises. Figure 1 shows Big data applications in various fields.

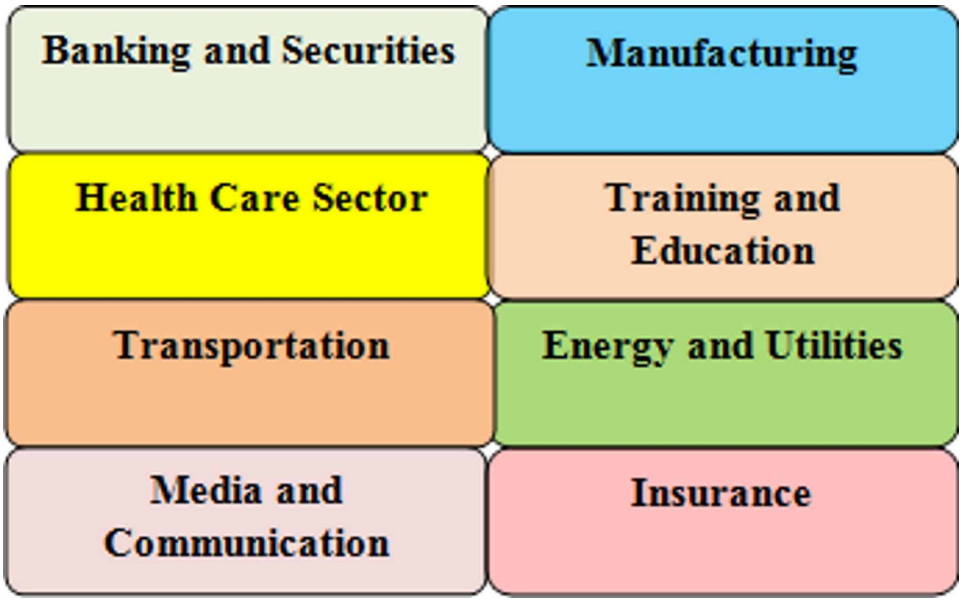
Aside from traditional patient data contained in content, there are different pictures and sounds recorded, from x-beams and ultrasounds, to Doppler and MRI imaging. A few specialists very much want that their discussions with patients be recorded for the patient's advantage (Qiu et al., 2015). This gathering of conflicting information is for the most part unstructured and can't be requested in the flawless tables and segments of a social database. This is the place big databases, as Hadoop, score. Be that as it may, it is one thing to store big data and very another to recover it seriously. Data researchers who can plan techniques to extricate important information from the non-consecutive and apparently irregular big databases are currently popular (Rath et al., 2018).

These techniques are challenging and hard to solve, however the IT business is beginning to convey activities that make significant data extraction less demanding. There is additionally a move to a half breed database structure, where data is put away in both a social and a "No SQL" database. Where social insurance elements have handled this obstacle, the outcome is an all-encompassing perspective of the patient, which expels a portion of the unpredictability of finding for the medicinal professional and makes life less difficult for the patient (Archenna et al., 2015). It likewise opens the path for the move to machine-to-machine (M2M) correspondence and the utilization of computerized reasoning to filter through and investigate data transmitted from the sensors gathering it (Rath et al., 2017). The future guarantee is examination that will screen health more than ever however there are likewise further issues that should be tended to, for example, data protection and security. Figure.1. Big data applications in various fields.

For analysis of major type of business process, Big Data helps in transforming major selling practice by improved techniques and accurate analysis of presented data. As shown in Figure 1 some of such process are Banking and Security, (Khan et al., 2014) Media and communication, Health Care Sector,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Big data applications in various fields



Manufacturing, Training and Education, Insurance, Transportation, Energy and utilities. Using big data analytics in all these fields improves the quality of the process and results improved output. The organisation of the paper has been done as follows. The Second section describes the Literature Review. Section 3 presents Big Data Applications in Modern Technology, Section 4 presents Monitoring and Controlling Health Care using Big Data Analytics and at last Section 5 concludes the paper.

2. LITERATURE REVIEW

Though there are many dangerous impacts of high radiations on health, radio signals on vision and many other such impact on health system, still the speed of progress in technology in social systems is impossibly fast. IBM Watson's electronic concept (Lomotey et al., 2017) accumulates atrocious measures of information, its impact and plans the best treatment elective for patients by checking terrifically vital medicinal examinations, and significant learning computations will do substantially more. Propelled tattoos can check the critical prosperity parameters and basic signs to disclose to us when there's short remark care of through our mobile phones (Vippalapalli et al., 2016). Expanded reality devices, for instance, Holocene from Microsoft can broaden propelled scenes on what we see and prepare masters for troublesome methodologies. Medicinal administrations advancement just doesn't stop stunning (Finlay et al., 2016). In 2013, there were bits of prattle about new organizations transforming adherence control by inserting microchips into calm compartments to allow specialists to check whether and when patients take them. Devices that allow holographic data input yet just misrepresented, toy-like contraptions have ended up being open. Without broad associations placing assets into them, they will have far to go before standard choice. A phenomenal case is L'Oréal, the excellence mind items association putting into a wearable sensor that fills the wearer in with respect to whether first experience with the Sun accomplishes dangerous levels (Pusala et al., 2016).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The present expansion and development in the field of Internet of Things (IoT) are giving surprising potential (Farahani et al., 2018) in the course of the novel era of healthcare. The visualization of the healthcare is expansively supported, as it progresses the perfection of life and health of people, including a few health controls. The perpetual increment of the multifaceted IoT devices in health is comprehensively tried by difficulties, for example, controlling the IoT terminal hubs utilized for health observing, continuous information handling and shrewd choice and occasion administration. In the choice of the most modern decade, there has been an expanding passion for huge information look into, particularly for health administrations applications (Rath et al., 2016). The requisition of the distributed computing and the Internet of Things (IoT) worldview in the healthcare field can convey a few chances to therapeutic IT, and specialists trust that it can essentially enhance healthcare administrations and add to its ceaseless and orderly advancement in a major information condition, for example, Industry 4.0 applications. In any case, the expected assets to oversee such information in a cloud-IoT condition are as yet a major test. In like manner, (Elhoseny et al., 2018) proposes another model to streamline virtual machines choice (VMs) in cloud-IoT health administrations applications to productively deal with a major measure of information in incorporated industry 4.0. Industry 4.0 applications require to process and examine huge information, which originate from various sources, for example, sensor information, without human mediation. The proposed demonstrate plans to improve the execution of the healthcare frameworks by lessening the partners' demand execution time, streamlining the required stockpiling of patients' huge information and giving an ongoing information recovery component for those applications.

Real time information can likewise be fused into big data. Ongoing big data has numerous preferences. For instance, any mistakes or investigate in an association can be recognized quickly and the operational issue can be survived. This will spare time, cost and increment the efficiency. The administrations additionally can be additionally enhanced as the continuous gives the most recent data on the topic. For example, it will give the total data on the patients and in the meantime ready to oversee therapeutic intercession immediately. In healthcare, big data is likewise utilized in prescient examination which is to distinguish and address the restorative issues previously it is turning into an unmanageable issue (Rath et al., 2014). Healthcare experts can lessen the hazard and beat the issue with the data got from the big data. Aside from that, big data is additionally ready to help recognize fakes in healthcare particularly on protection claims. Fake, irregularity and false claims can be hailed. This will encourage insurance agencies to avert misfortunes.

Big data can likewise profit healthcare through data administration, electronic therapeutic records and data examination. The big data will discover and distinguish the correct populace or target gathering. Big data comprises of assorted gathering of populace and certain gathering can be recognized for chance appraisal and screenings. The presence of big data will likewise permit advancement or change of a program or intercession to focus on the health issue (Khalifa et al., 2016). It will empower clinical preliminaries to be started instantly. Big data will give a clearer picture on the sort of populace and in addition their restorative issue. The example of the dispersion or malady data will permit fast advancement of mediation program and additionally focusing on the influenced bunch as ahead of schedule as conceivable Data development of pharmaceutical businesses were gotten from patients, caregivers, retailers and Research and Development.

The inclination of universal frameworks is supported by the improvement and dynamic selection of the Internet of Things (IoT) devices and their empowering advancements. IoT has been appeared to have critical potential in high-chance Environment, Health, and Safety (EHS) enterprises. In these businesses, human lives are in question and IoT-based applications are prepared to offer protected, dependable, and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

effective arrangements because of their capacity to work at a fine granular level and give rich low-level data. (Thibaud et al., 2018) audits existing distributed research on IoT-based applications in high-hazard EHS businesses with particular prominence on healthcare industry, nourishment inventory network (FSC), mining and vitality enterprises (oil and gas and atomic), insightful transportation (e.g., associated vehicles), and building and foundation administration for crisis reaction tasks until 2016.Web of Things (WoT) offers a consistent stage to interface individuals and items to each other for improving and making our lives simpler. This vision conveys us from form based unified plans to a more circulated condition offering a tremendous measure of utilizations, for example, intelligent wearable devices, smart home, perceptive versatility, and savvy urban areas. Farahani et al. (2018) talks about material ness of IoT in healthcare and prescription by displaying an all-encompassing design of IoT eHealth biological community. Healthcare is ending up progressively hard to oversee because of inadequate and less compelling healthcare administrations to meet the expanding requests of rising maturing populace with constant ailments. This requires a progress from the center driven treatment to understanding driven healthcare where every operator, for example, clinic, patient, and administrations are flawlessly associated with each other. This patient-driven IoT eHealth biological community needs a multi-layer design: (1) device, (2) mist figuring and (3) cloud to engage treatment of complex information regarding its assortment, speed, and inertness. This mist driven IoT design is trailed by different case cases of administrations and applications that are actualized on those layers. Those illustrations run from versatile health, helped living, e-drug, inserts, early cautioning frameworks, to populace checking in keen urban communities.

Over the most recent couple of years, the m-health care (mobile health care) applications in context of Internet of Things (IoT) have given multi-dimensional highlights and ongoing administrations. These applications give a stage to a huge number of individuals to get health refreshes routinely for a healthier way of life. Enlistment of IoT devices in the healthcare condition have renewed numerous highlights of these applications. The huge information created by IoT devices (Rath et al., 2015) in healthcare space is broke down on the cloud rather than exclusively depending on restricted capacity and calculation assets of handheld devices. In respect to this unique situation, a cloud-driven IoT based M-healthcare checking illness diagnosing system is proposed which predicts the potential ailment with its level of seriousness. Key wordings are characterized to produce client situated health estimations by investigating the idea of computational sciences. The structural model for brilliant understudy healthcare is intended for application situation. The outcomes are registered subsequent to preparing the health estimations in a particular setting. (Verma et al., 2018) deliberate understudy point of view health information is created utilizing UCI dataset and medicinal sensors to anticipate the understudy with various illness seriousness. Finding plans are connected utilizing different best in class order calculations and the outcomes are registered in light of precision, affectability, specificity, and F-measure. Exploratory outcomes demonstrate that the proposed technique beats the benchmark strategies for illness forecast.

3. BIG DATA ANALYTICS IN HEALTH CARE USING MODERN TECHNOLOGY

With the assistance of big data, the immense measure of data can be put away deliberately. Presently specialists and other human services professionals can settle on educated choices as they approach an extensive variety of data. Obviously, the data produced will develop significantly, and more up to date frameworks will have the capacity to process it rapidly and cost adequately. Tumor is quickly devastating people over the world. Big data can battle disease all the more successfully. Social insurance suppliers

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

will have upgraded capacity to recognize and analyze illnesses in their beginning periods, doling out more efficacious treatments in view of a patient's hereditary cosmetics, and manage sedate measurements to limit symptoms and enhance adequacy. It will likewise give awesome help to parallelization and help in mapping the 3 billion DNA base sets. The utilization of big data makes it less demanding for healing facility staff to work all the more effectively. Sensors are utilized other than understanding beds to ceaselessly screen pulse, pulse and respiratory rate. Any adjustment in design is immediately alarmed to specialists and social insurance chairmen.

New prospects for health care checking is induced by the expansion of Internet of Things (IoT) and Big data, and in addition the universal idea of little wearable bio sensors (Jeste et al., 2016) Numerous defy presently can't seem to be routed to make a tried and true and adaptable framework for health care checking Big data investigation on IoT based Health care framework is proposed and it utilizes IoT based health care observing framework that contains "Web of health sensor things". These things create gigantic volumes of data that couldn't be dealt with by the doctor. The Physicians imperative concern is that they have to settle on basic choices about their patient's health from these gigantic volumes of health data. He/she needs to isolate the data around one specific patient from the surge of health care data landing from the monstrous number of patients. Intel Galileo Gen 2 is going about as IoT operator and is utilized to convey the health data of patients into the Cloud. The could deal with the expanding volume of health data, ingeniously share the data crosswise over healthcare frameworks and give feasibility to Big data examination. Ongoing alarming of patient health data is a critical exercise in Big data which is essential in the proposed work. The Big health sensor data are investigated utilizing the Hadoop system.

Big data could encourage the pharmaceutical organizations to recognize new potential and viable medications and convey it to the clients all the more rapidly. Issues with Big Data, there is an enormous test in big data as far as data assurance, gathering and sharing of health data and data utilization. Big data examination with the utilization of refined innovations can possibly change the data storehouses and settle on educated choices. Issues, for example, protection, security, gauges and administration to be tended to. Data, for example, nano particulate treatment on growth treatment could be likewise be fused in big data to give a diagram and best treatment for disease particularly when nanotechnology is imperative in tranquilize conveyance in tumour treatment. Security Since the big data contained subject's close to home data and their health history, it is critical for the database to be shielded from hacking, digital burglary and phishing, where the stolen data can be sold for a colossal aggregate. Aside from the health data and individual data from the health framework which can be hacked or stolen, other big data in other business associations, for example, broadcast communications companies, banks or monetary organization are additionally powerless without the information of the customers. Before big data can be executed, it is important to guarantee that the organization, security, security of the big data (Rath et al., 2018) is all around ensured. Insurance health data by means of transmission security, multilayer confirmation, utilizing against infection programming, firewalls, data encryption are without a doubt fundamental. As the data turns out to be more local and worldwide, it turns out to be more convoluted and have more genuine effect on security, guidelines, dialect and phrasing (Rath et al., 2018). The availability of the healthcare data should be reliably inspected and observed. Data Classification Big data is a monstrous, less organized and heterogeneous. In that capacity there is a need to distinguish and order the data with the goal that it very well may be utilized successfully. Nonetheless, it is relentless to look for a particular data in the big data. The big data additionally required to be contextualized or pooled together with the goal that it will turn out to be more important to particular people or gatherings. Data Modelling Although big data is astounding for demonstrating and simulation, there is a need to recognize,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

structure and pool the correct important data with the goal that it very well may be utilized to display the issues, which later can be utilized for intercession. Without the correct organized data, it is trying to break down and envision the yield and to extricate particular data or data.

The intensity of machine learning has been enormously perceived in changing different ventures and health care applications (Mertz et al., 2018). The reason that Amazon, Netflix, and Google have changed their enterprises and have inserted learning all through each part of what they do. AEGLE enterprise (Wang et al., 2017) focuses to construct an imaginative ICT arrangement tending to the entire data esteem chain for health in view of: distributed computing empowering dynamic asset portion, HPC frameworks for computational speeding up and propelled perception techniques. analysis of the tended to Big Data health situations and the key empowering advances, and data security and administrative issues has been completed which are to be incorporated into AEGLE's bio analytical community, empowering propelled health-care expository administrations, while additionally advancing related research exercises.

Basic Health Care IT Computing Systems such as life sustained devices, Health checking systems, data processing systems that give purpose of caring solution and direction to care groups are a key segment of a lifesaving exertion in Healthcare. The high trends of mobility, social, cloud joined with far reaching increment and modernity of malware, has made new difficulties and the point in time location strategies at the clinics are never again compelling and represent a major danger to the basic care systems. To keep up the accessibility and honesty of these basic care systems, new versatile, learning security safeguard systems are required that not just gains from the activity entering the healing facility, yet in addition proactively gains from the movement around the world. Cisco's Cloud web security (CWS) (Rogers et al., 2017) gives industry-driving security and control for the dispersed attempt by ensuring clients all over, whenever through Cisco overall danger knowledge, propelled risk barrier abilities, and wandering client assurance. It uses the huge information to perform conduct investigation, irregularity location, avoidance obstruction, quick Detection administrations utilizing stream based, signature based, conduct based, and full bundle catch models to distinguish dangers. A specialized audit takes a gander at how huge Data Analytics is utilized as a part of blend with other security capacities to proactively distinguish dangers and anticipate far reaching harm to social insurance basic resources.

3.1. Interfacing the Physical World to the Internet of Things (IoT)

It takes into account the advancement of a wide assortment of utilizations. Things can be sought, overseen, examined, and even incorporated into collective amusements. Enterprises, health care, and urban areas are misusing IoT information driven systems to make these associations more productive, in this way, enhancing the lives of residents. For making IoT a reality, information created by sensors, advanced mobile phones, watches, and different wearable should be coordinated; additionally, the importance of IoT information ought to be expressly spoken to. Nonetheless, the Big Data nature of IoT information forces challenges that should be tended to keep in care the end goal to give versatile and productive IoT information driven foundations. These issues have been taken care of and focussed on the problems of portraying the significance of IoT spilling information utilizing philosophy and incorporating this information in a learning chart (Hossain et al., 2017). DESERT, a SPARQL inquiry motor has been planned to influence ready to on-Demand factoriz and Semantically To advance stReam information (DESERT) in a learning diagram. Coming about learning charts show the semantics or significance of combined information as far as elements that fulfill the SPARQL inquiries and connections among those substances; in this manner, just information required for question noting is incorporated into the learning diagram. It

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

assesses the consequences of DESERT on SRBench a benchmark of Streaming RDF information. The trial comes about recommend that DESERT takes into consideration accelerating question execution while the extent of the learning charts remains generally low.

One of the biggest obstacles obstructing to utilize big data in social insurance is the manner by which restorative data is spread crosswise over numerous sources represented by various states, clinics, and authoritative offices. Combination of these data sources (Panda et al., 2017) would require building up another foundation where all data suppliers team up with each other. Similarly, critical is actualizing new data examination instruments and procedures. Social insurance needs to get up to speed with different ventures that have effectively moved from standard relapse-based techniques to more future-arranged like prescient investigation, machine learning, and chart examination.

3.2. System Used for Alert in Real Time Applications

Data security and patient secrecy is an additional hazard. One of the benefits of paper and siloed electronic records is that they are not hackable. At the point when human services associations consider versatile application advancement to help IoT equipment, they should likewise guarantee that weakness is held within proper limits. Regardless of these dangers, wearable, sensors and portable applications are propelling the wildernesses of medicinal learning at a consistently expanding rate – particularly in battling and anticipating perilous sicknesses and immune system diseases. The odds of creating tumor in one's lifetime are evaluated to be as high as one out of two (for US guys). In any case, the industry is making fast walks in battling and keeping these malignancies. Different cases of big data in human services share one pivotal usefulness – constant alarming. In doctor's facilities, Clinical Decision Support (CDS) programming breaks down restorative data on the spot, furnishing health experts with guidance as they settle on prescriptive choices. Be that as it may, specialists need patients to avoid healing centers to maintain a strategic distance from expensive in-house medicines. Individual examination equipment, as of now drifting as business intelligence popular expressions a year ago, can possibly turn out to be a piece of another medicinal services.

3.3. Electronic Health Records (EHRs)

It's the most across the board use of big data in human services. Each patient has his own advanced record which incorporates socioeconomics, restorative history, hyper sensitivities, research facility test comes about and so on. Records are shared by means of secure information frameworks and are accessible for social insurance suppliers from both open and private area. Each record is included one modifiable document, which implies that specialists can execute changes after some time with no printed material and no peril of data replication.

EHRs can likewise trigger admonitions and updates when a patient ought to get another lab test or track medicines to check whether a patient has been following specialists' requests. In spite of the fact that EHR are an awesome thought, numerous nations still battle to completely execute them. U.S. has made a noteworthy jump with 94% of healing centers receiving EHRs as indicated by this HITECH research, yet the EU still lingers behind. Notwithstanding, an aggressive mandate drafted by European Commission should transform it: by 2020 brought together European health record framework ought to end up a reality.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As a case study, Kaiser Permanente is driving the path in the U.S. and could give a model to the EU to take after. They've completely executed a framework called Health Connect that offers data over the majority of their offices and makes it less demanding to utilize EHRs. A McKinsey cover big data in human services expresses that "The incorporated framework has enhanced results in cardiovascular infection and accomplished an expected $1 billion in investment funds from lessened office visits and lab tests." In any case, there are some sublime occurrences where medicinal services doesn't fall behind, for example, EHRs (particularly in the US). So, regardless of whether these administrations are not some tea, you are a potential patient, thus you should think about new social insurance (Patel et al., 2017) investigation applications. In addition, it's great to investigate now and again and perceive how different ventures adapt to big data.

One of the key data sets is 10 years of clinic confirmation records, which data researchers crunched utilizing "time arrangement examination" systems. These examinations enabled the researchers to see significant examples in confirmation rates. At that point, they could utilize machine figuring out how to locate the most exact calculations that anticipated future affirmations patterns. Summing up the result of this work, Forbes states: "The outcome is an internet browser-based interface intended to be utilized by specialists, attendants and clinic organization staff – untrained in data science – to figure visit and confirmation rates for the following 15 days. Additional staff can be drafted in when high quantities of guests are normal, prompting decreased sitting tight circumstances for patients and better nature of care."

3.4. The Utilization of Sensors to Monitor Activities

One of the key components adding to this war on tumor is the capacity to accumulate large volumes of patient data amid clinical preliminaries, utilizing sensors as opposed to overviews and meetings. Patients with different myeloma, an exceptionally agonizing blood growth, are to be observed through a blend of wearable and cell phones, with data transferred to a cloud stage. These patients are not being observed for their reaction to drugs, but instead for their personal satisfaction. Their rest examples and general portability are estimated to show signs of improvement comprehension of how the illness and ensuing treatments affect their everyday lives. Such research can likewise be utilized to advise and help health experts. Specialists of today need to remain over improvements in the adequacy of medications. A stage that both orders oncology research and keeps specialists educated enhances malignancy treatment on two fronts. Medicinal services are a segment in which IoT and big data will assume an undeniably fundamental part. While big data isn't at present required from a volume or speed point of view, this will change when data from billions of sensors should be put away and investigated. Table 1. Shows the description of some important contribution of Big Data associated with other technology in smart applications of society that are studied and analysed in this research work.

3.5. Precision of Medication

The Precision Medicine Initiative calls for restorative specialists to apply research and brought together data to advance customized quiet care. President Obama has a dream to set up a national patient databank to encourage customized treatment and advance genome research. Indeed, even without a national databank, healing centers can apply similar standards to give customized mind utilizing big data investigation to associate patient data put away in EMRs to national patterns and other data sources. Prescient investigation, for instance, can advance precaution nurture coronary illness and heftiness.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Description of some important contribution of Big Data associated with other technology

Literature / Author	Year	Details of Related Application
F. Zuhra et al.	2017	Use of WBSN (Wireless Body sensor Network and Protocols in detection of body anomalies
B. Farahani et al.	2018	Fog Computing approach and IoT based e-Health care system
M. Hossain et al.	2017	Internet of Things (IoT) Based Health Prescription Assistance software for health diagnosis
A. Harbouche et al.	2017	Use of WBN (Wireless Body Sensor Network) for health care and correct monitoring system
T. Adame et al.	2018	A Hybrid monitoring system for smart health care using RFID and Wireless Sensor Network
J. Lloret et al.	2017	Improved Protocol for efficient and smart nonstop e-health controlling & monitoring
O. Olayinka et al.	2017	Big data knowledge Implementation in Global Health Education System
B. Knoppers et al.	2017	Rules related to Ethics and Systems in big data health care Analysis
M. Pramaink et al.	2017	Big data enabled health paradigm and issues in smart city Health Applications

3.6. Proof Based Prescription

Cookbook drug has been the standard, utilizing a similar battery of tests to analyze by decision out the reason for ailment. With confirm based prescription, specialists connect side effects to limit the determinations. Beth Israel Deaconess Medical Center in Boston, for instance, is utilizing persistent data from two million patients to give data focuses to diagnosing through a cell phone application. Keeping in mind the end goal to encourage investigation, doctor notes are being encoded to institutionalize references; for instance, "hypertension" and "raised circulatory strain" are coded similarly to make data accessible.

Perceptive investigation likewise advances quality care and patient wellbeing. In the emergency unit, illustration, patients are inclined to a sudden downturn because of sepsis or different diseases. Sepsis alone has a 40 percent death rate and is hard to distinguish for early treatment. The University of California, Davis, has utilized EMR data investigation to make a calculation to give an early cautioning of sepsis contamination. In another illustration, the University of Iowa Hospitals and Clinics has utilized prescient examination to diminish postoperative disease following colon medical procedure by 58 percent. Similarly, as examination can be utilized for prescient administer to singular patients, similar techniques can be utilized as a part of the study of disease transmission. Duke University, for instance, is mapping EMRs to geographic information framework data so as to distinguish human services slants in particular geographic territories. By mapping the correct data sets, it's conceivable to anticipate particular illnesses, for example, flu that will raise in particular regions, making it less demanding to strategize diagnostics and plan for stocking serums and antibodies. The National Institutes of Health has been taking a shot at comparable systems for quite a while to anticipate malady episodes.

3.7. Reduced Rate of Hospital Re-Admissions

Readmission rates are a ceaseless issue for doctor's facilities, particularly with patients who return inside 30 long periods of treatment. Big data investigation of EMRs uncovers slants that feature those patients prone to require extra treatment to forestall readmission. At a Texas healing center, for instance, EMR investigation prompted a drop in the readmission rate of cardiovascular patients from 26.2 percent to 21.2 percent by recognizing high-chance patients. Commentators stress that patient records are a practical objective for digital hoodlums, on the grounds that they yield individual information that is significantly

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

more important than charge card data, for example, Social Security numbers, Medicare information and solution information. Advocates call attention to that big data investigation can really be a significant resource in anchoring therapeutic records by distinguishing changes in arrange activity or conduct that demonstrates a digital assault. Restorative protection is mind boggling and tormented with questioned and fake cases. Big data examination streamlines the efficiency of medicinal protection guarantees by uncovering claim slants and streamlining claims preparing. Patients show signs of improvement returns on their protection cases, and parental figures get instalment speedier.

These are only a couple of the ways that big data investigation is having an effect in social insurance. Anyplace there is a data-driven process, big data investigation can be connected to enhance persistent care and health focus tasks. Furthermore, new innovation, for example, radio frequency recognizable proof (RFID) chips—is conveying more data for potential examination. New data sources give new chances to big data investigation, and it's up to arrangement suppliers to demonstrate to human services experts generally accepted methods to open the intensity of that data with new examination.

3.8. Customary Patient Checking Using IoT

The Internet-of-Things (IoT) has supposed control over the business range, and its applications fluctuate broadly from farming and social insurance to transportation. A hospital condition can be extremely upsetting, particularly for senior people and youngsters. With the consistently expanding total populace, the regular patient-physical checkup has lost its viability. Thus, keen medicinal services turn out to be critical. Savvy social insurance can be actualized at all levels, beginning from temperature checking for children to following essential signs in the elderly (Sunderavadivel et al., 2018). The intricacy and cost of execution shifts in light of the required exactness of the individual equipment, functionalities, and complexity of the application for which they are utilized. Savvy medicinal services likewise fall under vertical territories, for example, expansive scale combination, inserted framework, enormous information, machine learning, distributed computing, and computerized reasoning.

Data technology has progressed amid the most recent five decades to the phase where its effect is being felt by the general public in each administration that it gets from media, business, health care, shopper devices, vitality and power, and transportation domains. Amid this course of human-technology collaboration tremendous measure of data and learning exchange happens straightforwardly between specialist co-ops and their customers, and in addition by implication between customers. Since human propensity is to "investigate" its past with a specific end goal to foresee the "future", monitoring this powerfully gushing voluminous heterogeneous data, called Big Data (BD), and examining it for significant revelation of learning that prompts esteem included business turns into a critical research action. It is in this setting research in Big Data (BD) figuring has developed. Important choices can be construct just with respect to huge information disclosure, which thus requires a decent comprehension of the attributes of the aggregated data, a fitting grouping of this enormous gathering, and a proficient investigation of it. Health care part is a basic framework since its administrations influence the lives of humans and the absence of administration coherence might be unfortunate to the economy and human lives. The huge measure of data gathered by this part from its customers is organized into Electronic Health Records (EHR) which is BD and is utilized alongside pharmaceutical and administrative data in giving health administrations. More BD is produced while controlling administrations and estimating their effects on customers subsequent to regulating the administrations. Literature (Pusala et al., 2016) explores the sorts and features of Health Care BD (HBD), its attributes, and give a characterization of it.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4. MONITORING AND CONTROLLING HEALTH CARE USING BIG DATA ANALYTICS

Indeed, in the past plan, medicinal services suppliers had no immediate motivating force to impart persistent information to each other, which had made it harder to use the intensity of big data. Since a greater amount of them are getting paid in light of patient results, they have a money related motivation to share data that can be utilized to enhance the lives of patients while cutting expenses for insurance agencies. At long last, doctor choices are ending up increasingly confirm based, implying that they depend on large swathes of research and clinical data instead of exclusively their tutoring and expert feeling. As in numerous different enterprises, data social occasion and administration is getting bigger, and experts require help in the issue. This new treatment state of mind implies there is a more noteworthy interest for big data investigation in human services offices than any time in recent memory, and the ascent of SaaS business intelligence apparatuses is additionally noting that need.

4.1. Electronic Health Records

Doctors sharing EHRs (Electronic Health Records) can total and break down data for patterns that can decrease human services costs. Sharing data amongst doctors and social insurance suppliers as they inspect patients can decrease copy tests and enhance understanding consideration. Most EHR data is siloed, to a great extent for security and administrative consistence reasons, however finding a protected path to mine patient data can enhance the nature of care while lessening costs.

Consider the intensity of examining patterns in doctor's facility mind. For instance, concentrating investigation of medicinal instruments in a pediatric ward can detach conceivable newborn child disease drifts prior. Or then again consider the instance of one doctor's facility that could diminish post-agent staph contamination: by utilizing big data, the organization could establish that one specialist was recommending post-agent anti-microbials that were less powerful against disease. Table 2 displays the descriptions of various research related to big data in health care.

4.2. Public Health Research With Controlled Information

The therapeutic calling is suffocating in data. Therapeutic workplaces and healing centers submit data about restorative conditions and inoculations, however without big data, those data are good for nothing. Utilizing investigation standardizes crude patient data to fill holes in general health records that can influence controls and in addition giving better care. Utilizing proof-based medication, the specialist can coordinate side effects to a bigger patient database so as to go to a precise finding quicker and all the more effectively. Where big data assumes a part is absorbing information from various sources and normalizing the data, so one record that portrays "hypertension" maps to another that depicts "elevated blood pressure."

4.3. Reducing Hospital Re-Admit With Better Medicinal Practice

Hospital costs are rising very high somewhat due to high readmission rates inside 30 long stretches of patient discharge. Utilizing big data examination so as to distinguish in danger patients in light of previous history, outline information, and patient patterns, healing centers can recognize in danger patients

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Descriptions of various research related to big data in health care and other related applications

Literature	Year	Research Related to Big Data
L. Cui et al.	2016	Big data use for SDN and SDN based Big data
Y. Demchenko et al.	2014	Online education instructional model using big data curriculum
M. Khan et al.	2014	Value extraction and understanding of big data
M. Pusala et al.	2016	Data Analytic Tools and Challenges
Shah et al.	2016	Remote health care: cyber system
K. Lee et al.	2016	Map reduce based output analysis tool in Big data
T. White et al.	2012	Big data analysis and focus on data mining techniques
M. Lemoudden et al.	2015	Cloud generated logs using big data technology
D. Singh et al.	2015	Survey platform for big data analysis
A. Ayed et al.	2015	Map Reduce based text detection in Big data
N. Zhu et al.	2014	Cost efficient Mapreduce for Hadoop Clusters
K. Shvachko et al.	2010	Hadoop Distributed File System for Big Data Analytics
A. Sahni et al.	2015	Bandwidth analysis in cellular network using Big data analytics
J. Liu et al.	2014	Big traffic data analysis of a large-scale cellular network with Hadoop
A. Omar et al.	2015	Efficiency of cache nodes in CRN using Big data analysis
K. Zheng et al.	2016	Big data driven optimization for mobile networks in 5G
K. Pedersen et al.	2013	Inter cell Interference in LTE Network
C. Lee et al.	2014	Handover optimization in using big data analytics

and give the essential care to lessen readmission rates. Insurers like UnitedHealthcare are utilizing big data examination keeping in mind the end goal to recognize therapeutic extortion and fraud. The organization utilizes investigation on discourse to-content records from calls to the call focus to recognize potential fraudsters. The insurance agency additionally utilizes big data keeping in mind the end goal to anticipate which kinds of treatment designs will probably succeed.

As medicinal practices develop, obliging more specialists and more patients turns out to be additionally testing. Consider Westmed Medical Group in Westchester County, New York. This training developed from 16 doctors in 1996 to 250 doctors today observing 250,000 patients, with yearly income of $285 million. As the training develops, it should be more effective keeping in mind the end goal to succeed. Utilizing big data, the training could break down in excess of 2,200 procedures and methodology. Accordingly, the training could streamline work process, move clinical undertakings from specialists to medical caretakers, lessen pointless testing, and enhance persistent fulfillment. Like any business, big data made it obvious where procedures could be progressed.

The above issues are not many that are drastically affecting medicinal services. The more data that are accessible to doctors and clinic overseers, the less demanding it is to distinguish patterns, standardize persistent data, and recognize bottlenecks in understanding consideration. Restorative experts can apply big data investigation similarly they are connected in some other business. The main distinction is that the stakes are higher and the bits of knowledge from big data may enable spare to people's lives. For those with unending medical problems or perilous ailments, the line of pros and methodology can appear to be perpetual. Besides, each time another specialist or facility is gone to, the patient is required

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to fill in long and convoluted structures – particularly if a medicinal protection is included. It's presently evident that the business needs to experience some significant changes, beginning with data administration. In the course of the most recent decade, clinics, medicinal practices, therapeutic plans and research offices have been digitizing their data with an end goal to distribution center it successfully. They are presently beginning to swing to counselling firms to embrace big data as an answer for data stockpiling.

As it were, medicinal services costs are considerably higher than they ought to be, and they have been ascending for as long as 20 years. Plainly, we need some brilliant, data-driven reasoning here. What's more, current motivating forces are changing too: numerous insurance agencies are changing from charge for-benefit designs (which compensate utilizing costly and, in some cases, superfluous medicines and treating large measures of patients rapidly) to plans that organize persistent results. As the writers of the famous Freakonomics books have contended, money related impetuses matter – and motivating forces that organize patient's health over treating large measures of patients are something to be thankful for.

5. CONCLUSION

The influence of health care administrations using big data has been developing significantly faster than many similar technologies. progressions in health care informatics, digitalizing health records, and telemedicine has brought about fast development of health care data science. These days, there is a consistently relocation of people to urban zones. Health care benefit is a standout amongst the most difficult perspectives that is incredibly influenced by the huge flood of individuals to downtown areas. Thus, urban areas around the globe are putting vigorously in advanced change with an end goal to give healthier bio analytical systems to individuals. A standout amongst the most encouraging regions where big data can be connected to roll out an improvement is medicinal services. Medicinal services investigation can possibly decrease expenses of treatment, foresee flare-ups of scourges, keep away from preventable ailments and enhance the personal satisfaction all in all. Normal human life expectancy is expanding along total populace, which postures new difficulties to the present treatment conveyance techniques. Medicinal services experts, much the same as business people, are now equipped with reservoir of data and they can search for best procedures to utilize these useful datasets. The above article addresses the applications of big data in human service sector significantly in health care with an analytical review that includes role of big data in human service and medicinal sector.

REFERENCES

Adame, T., Bel, A., Carreras, A., Melià-Seguí, J., Oliver, M., & Pous, R. (2018). CUIDATS: An RFID–WSN hybrid monitoring system for smart health care environments. *Future Generation Computer Systems*, *78*, 602–615.

Archenaa, J., & Mary Anita, E. A. (2015). A Survey of Big Data Analytics in Healthcare and Government. *Procedia Computer Science*, *50*, 408–413. doi:10.1016/j.procs.2015.04.021

Cui, L., Yu, F. R., & Yan, Q. (2016). When big data meets software-defined networking: SDN for big data and big data for SDN. *IEEE Network*, *30*(1), 58–65.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Demchenko, Y., Gruengard, E., & Klous, S. (2014). Instructional model for building effective big data curricula for online and campus education. In *IEEE 6th International Conference on Cloud Computing Technology and Science* (pp. 935–941). IEEE.

Elhoseny, M., Abdelaziz, A., Salama, A. S., Riad, A. M., Muhammad, K., & Sangaiah, A. K. (2018). A hybrid model of internet of things and cloud computing to manage big data in health services applications. *Future Generation Computer Systems*, *86*, 1383–1394. doi:10.1016/j.future.2018.03.005

Farahani, B., Firouzi, F., Chang, V., Badaroglu, M., Constant, N., & Mankodiya, K. (2018). Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare. *Future Generation Computer Systems*, *78*(Part 2), 659–676. doi:10.1016/j.future.2017.04.036

Finlay, D. (2016). Article Seven - Connected Health Approaches to Wound Monitoring. In J. Davis, A. McLister, J. Cundell, & D. Finlay (Eds.), *Smart Bandage Technologies* (pp. 229–244). Academic Press. doi:10.1016/B978-0-12-803762-1.00007-2

Harbouche, A., Djedi, N., Erradi, M., Ben-Othman, J., & Kobbane, A. (2017). Model driven flexible design of a wireless body sensor network for health monitoring. *Computer Networks*, *129*(Part 2), 548–571. doi:10.1016/j.comnet.2017.06.014

Hussain, A., Wenbi, R., da Silva, A. L., Nadher, M., & Mudhish, M. (2015). Health and emergency-care platform for the elderly and disabled people in the Smart City. *Journal of Systems and Software*, *110*, 253–263. doi:10.1016/j.jss.2015.08.041

Hossain, M., Islam, S. R., Ali, F., Kwak, K. S., & Hasan, R. (2018). An Internet of Things-based health prescription assistant and its security system design. *Future Generation Computer Systems*, *82*, 422–439.

Hossain, M. S., Rahman, M. A., & Muhammad, G. (2017). Towards energy-aware cloud-oriented cyber-physical therapy system. *Future Generation Computer Systems*.

Jeste, D. V., Blazer, D. G. II, Buckwalter, K. C., Cassidy, K. L. K., Fishman, L., Gwyther, L. P., ... Vega, W. A. (2016). Age-friendly communities initiative: Public health approach to promoting successful aging. *The American Journal of Geriatric Psychiatry*, *24*(12), 1158–1170.

Khalifa, M., & Zabani, I. (2016). Utilizing health analytics in improving the performance of healthcare services: A case study on a tertiary care hospital. *Journal of Infection and Public Health*, *9*(6), 757-765.

Knoppers, B. M., & Thorogood, A. M. (2017). Ethics and big data in health. *Current Opinion in Systems Biology*, *4*, 53–57. doi:10.1016/j.coisb.2017.07.001

Lloret, J., Parra, L., Taha, M., & Tomás, J. (2017). An architecture and protocol for smart continuous eHealth monitoring using 5G. *Computer Networks*, *129*(Part 2), 340–351. doi:10.1016/j.comnet.2017.05.018

Lomotey, R. K., Pry, J., & Sriramoju, S. (2017). Wearable IoT data stream traceability in a distributed health information system. *Pervasive and Mobile Computing*, *40*, 692–707.

Mamata, R. (2018), An Analytical Study of Security and Challenging Issues in Social Networking as an Emerging Connected Technology. In *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies*, Malaviya National Institute of Technology, Jaipur, India, March 26-27. Elsevier.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mamata, R. B. P. (2018). Communication Improvement and Traffic Control Based on V2I in Smart City Framework. *International Journal of Vehicular Telematics and Infotainment Systems*, *2*(1).

Mertz, L. (2018). Machine learning takes on health care: Leonard D'Avolio's Cyft Employs Big Data to benefit patients and providers. *IEEE Pulse*, *9*(1), 10–11. doi:10.1109/MPUL.2017.2772686 PMID:29373850

Olayinka, O., Kekeh, M., Sheth-Chandra, M., & Akpinar-Elci, M. (2017). Big Data Knowledge in Global Health Education. *Annals of Global Health*.

Panda, M., Ali, S. M., & Panda, S. K. (2017). Big data in health care: A mobile based solution. In *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, Chirala (pp. 149-152). 10.1109/ICBDACI.2017.8070826

Patel, P., Intizar Ali, M., & Sheth, A. (2017). On Using the Intelligent Edge for IoT Analytics. *IEEE Intelligent Systems*, *32*(5), 64–69. doi:10.1109/MIS.2017.3711653

Pattanayak, B., & Rath, M. (2014). A mobile agent based intrusion detection system architecture for mobile ad hoc networks. *Journal of Computational Science*, *10*(6), 970–975. doi:10.3844/jcssp.2014.970.975

Pramanik, M. I., Lau, R. Y. K., Demirkan, H., & Azad, M. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems with Applications*, *87*, 370–383. doi:10.1016/j.eswa.2017.06.027

Pusala, M. K., Salehi, M. A., Katukuri, J. R., Xie, Y., & Raghavan, V. (2016). Massive Data Analysis: Tasks, Tools, Applications, and Challenges. In *Big Data Analytics* (pp. 11–40). New Delhi: Springer.

Qiu, H. J., Ho, I. W. H., Chi, K. T., & Xie, Y. (2015). A Methodology for Studying 802.11p VANET broadcasting performance with practical vehicle distribution. *IEEE Transactions on Vehicular Technology*, *64*(10), 4756–4769. doi:10.1109/TVT.2014.2367037

Rath, M. (2017). Resource provision and QoS support with added security for client side applications in cloud computing. *International Journal of Information Technology*, *9*(3), 1–8.

Rath, M. (2018, June). Smart Traffic Management System for Traffic Control using Automated Mechanical and Electronic Devices. [). IOP Publishing.]. *IOP Conference Series. Materials Science and Engineering*, *377*(1), 012201.

Rath, M. (2018). Effective routing in mobile ad-hoc networks with power and end-to-end delay optimization: well matched with modern digital IoT technology. In *Advances in Data Communications and Networking for Digital Business Transformation* (pp. 206–243). Hershey, PA: IGI Global; doi:10.4018/978-1-5225-5323-6.ch007

Rath, M. (2018). A methodical analysis of application of emerging ubiquitous computing technology with fog computing and IoT in diversified fields and challenges of Cloud Computing. *International Journal of Information Communication Technologies and Human Development*, *10*(2), 15–27. doi:10.4018/IJICTHD.2018040102

Rath, M. (2018). A Methodical Analysis of Application of Emerging Ubiquitous Computing Technology With Fog Computing and IoT in Diversified Fields and Challenges of Cloud Computing. *International Journal of Information Communication Technologies and Human Development*, *10*(2).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rath, M., & Misra, R. (2018). An exhaustive study and analysis of assorted application and challenges in fog computing and emerging ubiquitous computing technology. *International Journal of Applied Evolutionary Computation*, *9*(2), 17–32. doi:10.4018/IJAEC.2018040102

Rath, M., & Oreku, G. S. (2018). Security Issues in Mobile Devices and Mobile Adhoc Networks. In *Mobile Technologies and Socio-Economic Development in Emerging Nations* (pp. 180–202). Hershey, PA: IGI Global; doi:10.4018/978-1-5225-4029-8.ch009

Rath, M., & Panda, M. R. (2017). MAQ system development in mobile ad-hoc networks using mobile agents. In *IEEE 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, Noida, Japan (pp. 794-798).

Rath, M., & Panigrahi, C. (2016). Prioritization of Security Measures at the Junction of MANET and IoT. In *Second International Conference on Information and Communication Technology for Competitive Strategies*. ACM Publication. 10.1145/2905055.2905187

Rath, M., Pati, B., & Pattanayak, B. (2015). Energy Competent Routing Protocol Design in MANET with Real time Application Provision. *International Journal of Business Data Communications and Networking*, *11*(1), 50–60. doi:10.4018/IJBDCN.2015010105

Rath, M., Pati, B., & Pattanayak, B. (2015). Delay and power based network assessment of network layer protocols in MANET. In *2015 International Conference on Control, Instrumentation, Communication and Computational Technologies*, Kumaracoil, India (pp. 682-686). 10.1109/ICCICCT.2015.7475365

Rath, M., Pati, B., & Pattanayak, B. (2016). Resource Reservation and Improved QoS for Real Time Applications in MANET. *Indian Journal of Science and Technology*, *9*(36). doi:10.17485/ijst/2016/v9i36/100910

Rath, M., Pati, B., & Pattanayak, B. (2016). *Comparative analysis of AODV routing protocols based on network performance parameters in Mobile Adhoc Networks. In Foundations and Frontiers in Computer, Communication and Electrical Engineering* (pp. 461–466). CRC Press, Taylor & Francis.

Rath, M., Pati B, Pattanayak B (2016), QoS Satisfaction in MANET Based Real Time Applications. *International Journal of Control Theory and Applications*, *9*(7), 3069-3083.

Rath, M., Pati, B., & Pattanayak, B. (2016). Energy Efficient MANET Protocol Using Cross Layer Design for Military Applications. *Defence Science Journal*, *66*(2), 146. doi:10.14429/dsj.66.9705

Rath, M., Pati, B., & Pattanayak, B. K. (2016). Inter-Layer Communication Based QoS Platform for Real Time Multimedia Applications in MANET. In Wireless Communications, Signal Processing and Networking (IEEE WiSPNET), Chennai, India (pp. 613-617). doi:10.1109/WiSPNET.2016.7566203

Rath, M., Pati, B., & Pattanayak, B. K. (2017). Cross layer based QoS platform for multimedia transmission in MANET. In *11th International Conference on Intelligent Systems and Control (ISCO)*, Coimbatore, Italy (pp. 402-407). 10.1109/ISCO.2017.7856026

Rath, M., Pati, B., & Pattanayak, B. K. (2018). Relevance of Soft Computing Techniques in the Significant Management of Wireless Sensor Networks. In Soft Computing in Wireless Sensor Networks (pp. 86-106). Chapman and Hall/CRC, Taylor & Francis Group. doi:10.1201/9780429438639-4

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rath, M., Pati, B., & Pattanayak, B. K. (2019). Mobile Agent-Based Improved Traffic Control System in VANET. In A. Krishna, K. Srikantaiah, & C. Naveena (Eds.), *Integrated Intelligent Computing, Communication and Security*. Singapore: Springer. doi:10.1007/978-981-10-8797-4_28

Rath, M., Pati, B., Pattanayak, B. K., Panigrahi, C. R., & Sarkar, J. L. (2017). Load balanced routing scheme for MANETs with power and delay optimisation. *International Journal of Communication Networks and Distributed Systems*, *19*(4), 394–405.

Rath, M., & Pattanayak, B. (2016). A Contemporary Survey and Analysis of Delay and Power Based Routing Protocols in MANET. *Journal of Engineering and Applied Sciences (Asian Research Publishing Network)*, *11*(1), 536–540.

Rath, M., & Pattanayak, B. (2017). MAQ: A Mobile Agent Based QoS Platform for MANETs. *International Journal of Business Data Communications and Networking*, *13*(1), 1–8. doi:10.4018/IJBDCN.2017010101

Rath, M., & Pattanayak, B. K. (2014). A methodical survey on real time applications in MANETS: Focussing On Key Issues. In *International Conference on, High Performance Computing and Applications (IEEE ICHPCA)*, Bhubaneswar, India (pp. 1-5, 22-24). 10.1109/ICHPCA.2014.7045301

Rath, M., & Pattanayak, B. K. (2017, December). SCICS: A Soft Computing Based Intelligent Communication System in VANET. In *International Conference on Intelligent Information Technologies* (pp. 255–261). Singapore: Springer; . doi:10.1007/978-981-10-7635-0_19

Rath, M., & Pattanayak, B. K. (2018). Monitoring of QoS in MANET Based Real Time Applications. In S. Satapathy & A. Joshi (Eds.), Information and Communication Technology for Intelligent Systems (Vol. 2, pp. 579-586). Cham: Springer. doi:10.1007/978-3-319-63645-0_64

Rath, M., & Pattanayak, B. K. (2018). Monitoring of QoS in MANET Based Real Time Applications. In S. Satapathy & A. Joshi (Eds.), Information and Communication Technology for Intelligent Systems (Vol. 2, pp. 579-586). Springer. doi:10.1007/978-3-319-63645-0_64

Rath, M., Rout, U. P., Pujari, N., Nanda, S. K., & Panda, S. P. (2017). Congestion Control Mechanism for Real Time Traffic in Mobile Adhoc Networks. In *Computer Communication, Networking and Internet Security* (pp. 149–156). Singapore: Springer.

Rath, M., Swain, J., Pati, B., & Pattanayak, B. K. (2018). Attacks and Control in MANET. In Handbook of Research on Network Forensics and Analysis Techniques (pp. 19–37). Hershey, PA: IGI Global.

Rath, M., Pattanayak, B. K., & Pati, B. (2017). Energetic Routing Protocol Design for Real-time Transmission in Mobile Ad hoc Network. In Computing and Network Sustainability (pp. 187-199). Springer, Singapore.

Rogers, W. A., & Mitzner, T. L. (2017). Envisioning the future for older adults: Autonomy, health, well-being, and social connectedness with technology support. *Futures*, *87*, 133–139.

Sahoo, J., & Rath, M. (2017). Study and Analysis of Smart Applications in Smart City Context. In *2017 International Conference on Information Technology (ICIT)*, Bhubaneswar, India (pp. 225-228). 10.1109/ICIT.2017.38

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shah, T., Yavari, A., Mitra, K., Saguna, S., Jayaraman, P. P., Rabhi, F., & Ranjan, R. (2016). Remote health care cyber-physical system: Quality of service (QoS) challenges and opportunities. *IET Cyber-Physical Systems: Theory & Applications*, *1*(1), 40–48.

Sun, E., Zhang, X., & Li, Z. (2012). The Internet of Things (IOT) and cloud computing (CC) based tailings dam monitoring and pre-alarm system in mines. *Safety Science*, *50*(4), 811–815. doi:10.1016/j.ssci.2011.08.028

Sundaravadivel, P., Kougianos, E., Mohanty, S. P., & Ganapathiraju, M. K. (2018). Everything You Wanted to Know about Smart Health Care: Evaluating the Different Technologies and Components of the Internet of Things for Better Health. *IEEE Consumer Electronics Magazine*, *7*(1), 18–28.

Thibaud, M., Chi, H., Zhou, W., & Piramuthu, S. (2018). Internet of Things (IoT) in high-risk Environment, Health and Safety (EHS) industries: A comprehensive review. *Decision Support Systems*, *108*, 79–95.

Uddin, M. F., & Gupta, N. (2014, April). Seven V's of Big Data understanding Big Data to extract value. In 2014 Zone 1 Conference of the American Society for Engineering Education (ASEE Zone 1) (pp. 1-5). IEEE.

Verma, P., & Sood, S. K. (2018). Cloud-centric IoT based disease diagnosis healthcare framework. *Journal of Parallel and Distributed Computing*, *116*, 27–38. doi:10.1016/j.jpdc.2017.11.018

Vippalapalli, V., & Ananthula, S. (2016). Internet of things (IoT) based smart health care system. In *International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, Paralakhemundi (pp. 1229-1233). 10.1109/SCOPES.2016.7955637

Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, *55*(1), 64–79.

Zuhra, F. T., Abu Bakar, K., Ahmed, A., & Tunio, M. A. (2017). Routing protocols in wireless body sensor networks: A comprehensive survey. *Journal of Network and Computer Applications*, *99*, 73-97.

This research was previously published in the International Journal of Big Data and Analytics in Healthcare (IJBDAH), 3(2); edited by Mu-Yen Chen; pages 45-61; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 93
The Impact of Healthcare Information Technology on Patient Outcomes

Edward T. Chen
University of Massachusetts Lowell, USA

ABSTRACT

The purpose of this article is to review the existing literature on health information technology (HIT), specifically electronic health records (EHR) and life-long patient records, in order to provide a broad assessment of the current state of this technology in the United States. Relevant literature was reviewed to determine whether key hypotheses were validated. Areas for additional research and development were identified, and a potential path forward was proposed. HIT adoption is a worthwhile effort in the United States, and it is possible for us to enact an interoperable central records system within our current fee-for-service healthcare system. Wide scale adoption will require subsidies and regulatory involvement at the state level, but professional networks may be exploited to speed the rate of adoption. A four-tier architecture with autonomic security systems, properly validated, can provide the infrastructure necessary.

INTRODUCTION

There can be no doubt that one's health is a major factor in quality of life. According to the World Health Organization (WHO), in their online World Bank Statistics, 9.7% of the world's total gross domestic (GDP) product was spent on healthcare in 2014. That equates to trillions of dollars spent annually on the direct delivery of healthcare from skilled health workers. This includes in-patient hospital care, out-patient preventative care, diagnosis and treatment, laboratory testing and imaging, emergency medical services and long-term care. The range of conditions is wide and disparate – everything from acute illness to chronic conditions, prenatal to end of life, physical and mental health. There is no facet of our life that does not, in some way, interact with the healthcare industry at one point or another.

DOI: 10.4018/978-1-7998-1204-3.ch093

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Yet, this important industry does not always deliver the standard of care it ought to. In the United States, annual spending exceeds $2.7 Trillion. This is the highest percentage of GDP for any nation on Earth, and yet our quality of care is ranked 37th in the world by the WHO. Clearly, there is a disconnection between our spending and outcomes (The World Bank, 2017).

The advancement of information technology in health care is pivotal to the guidelines set in the Patient Protection and Affordable Care Act. As health care organizations have become clinically data driven, the need for information systems and IT is imperative (Vo & Bhaskar, 2012). Analysts in IT departments at hospitals play the role in maintaining and implementing systems. Clinical staff members, such as nurses and physicians, play the role of end users and require the skills needed to effectively operate these systems and forms of technology in order to carry out their clinical duties (Najaftorkaman, Ghapanchi, Talaei-Khoei, & Ray, 2015). Key developments in IT include the transformation of the medical record from paper to digital, the use of mobile devices, and the emergence of integrated health care systems that have the capacity for handling the many complexities that the industry has to offer.

Equally impacted are all U.S. employers who offer health insurance to their employees and the dependents of their employees. While employers attempt to contain the trend of health insurance expense, information systems are needed to pinpoint conditions that prove to be the most costly and implement benefit redesign plans to reduce costs and promote wellness. Employers also keep their staff productive by keeping them healthy. In the same light, information systems help employees examine disease management programs that are geared towards steering employees to engage in healthy behavior and seek preventive care regularly. If health care was once described as being technically challenged when compared with other industries, that notion is being put to rest by the progression of IT in the industry (Vo & Bhaskar, 2012).

A potential bridge to creating a level of care commensurate with our spending is the use of integrated health information technology (HIT). It is theorized that increased use of HIT will reduce medical errors, streamline operations, and reduce overall costs while improving healthcare outcomes. In 2009, the US committed to a set of policies called the Health Information Technology for Economics and Clinical Health (HITECH) Act, aimed at spurring widespread use of Electronic Health Records (HER) and a level of interoperability among individual facilities and care providers (Adler-Milstein & Jha, 2014). Related to this Act, President Barak Obama committed $20 billion to computerizing medical records. Yet, adoption of EMR in the United States remains slow, with only 30% of hospitals across the country participating in the Health Information Exchange as of 2012.

This paper will explore the literature documenting the benefits of HIT. Case studies from other countries provide useful insights into the actual validity of theorized and purported benefits. We will explore two such cases. The risks of implementing HIT will be discussed, along with some of the mitigating approaches that can be used to reduce the risk. Finally, we will look more closely at the implementation barriers that have kept this technology from having more impact on our healthcare system.

STUDIES ON PATIENT OUTCOMES

The Patient Protection and Affordable Care Act (PPACA) was signed by President Barack Obama on March 3, 2010 and upheld by Congress on June 28, 2012. The act mandates that all Americans are required to be insured with health coverage or pay a penalty to the government beginning in 2014. Another

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

provision includes the expansion of Medicaid to adults whose incomes are at a defined rate at or below the federal poverty limit. Furthermore, the act seeks to decrease the cost of health care (Nussbaum, Tirrell, Wechsler, & Randall, 2010; Rosenbaum, & Margulies, 2011).

Exchanges have been established in the majority of the states either on their own or with the assistance of the Federal Government. The exchanges represent online marketplaces where participants can shop for health insurance plans. The exchanges also provide online resources that allow users to determine their eligibility for a federal subsidy (Carey & Appleby, 2013; Najaftorkaman, et al., 2015).

There are two primary claims regarding the potential benefits of implementing HIT. The first is that implementing HIT will increase the quality of outcomes in healthcare delivery. A major potential benefit is the reduced likelihood of healthcare mistakes, which can have severe and far-reaching impacts. According to Fichman, Kohli and Krishnan (2011), medical errors are a leading cause of adverse events that are expensive, increase patient stays, decrease patient outcomes, cost human lives, and can even affect public health. Aron et al. (2011) explored the potential relationship between HIT and procedural and interpretive medical errors across three primary healthcare processes – sensing, controlling and monitoring. They found that HIT could overcome interpretive errors when complemented by professional training that enables users to exploit automation. The impact on procedural errors was less clear. Although it was found that in some cases technology might increase the likelihood of those types of errors, a record of behavior could reinforce positive behavioral and procedural actions. This study also found that HIT can have a positive impact on public health outcomes, such as outbreaks of severe acute respiratory syndrome (SARS), when the technology has the ability to adapt to connect and disconnect key players as needed by the developing situation.

STUDIES ON FINANCIAL BENEFITS

The second proposed benefit is reduced medical spending. Section 3022 of the PPACA introduces a Medicare Shared Savings Program for the Accountable Care Organization (ACO). An Accountable Care Organization is a health care provider that aims at delivering high service quality care and reducing the cost of care. In doing so, the ACO can expect to achieve an incentive in the form of shared Medicare savings (Berwick, 2011).

The ACO model depends greatly on Health Information Technology (HIT). HIT has evolved in the health care industry more and more as data becomes the centerpiece of delivering high service quality care, cost analysis and payment reform. The ACO concept links care and payment risk across multiple organizations and physicians. In order for the model to survive, data sharing will be crucial. The delivery and storage of this data is where technical resources come into play. The ACO portrays a team concept where multiple entities collaborate to achieve financial incentive while providing the best possible care. In order to do so, a patient's Electronic Health Record (EHR) must be electronically maintained and available across all clinical professionals involved. The EHR concept is at the early stages of development in health care and currently represents limited data sharing capabilities. It is evident that Accountable Care Organizations will be required to invest in IT interfaces that track the patients across the entire episode of care (Bitton & Filer, 2012). Clinical data is at the forefront of the ACO model. As organizations adopt the concept, they will require sophisticated IT resources to share data internally and externally with other acute and post-acute care facilities.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

A study by Park et al. (2015) of a pilot HIT program in South Korea examined acceptance and use of HIT by health professionals, while also examining the outcomes in terms of financial and health impacts. This study followed a 3-year pilot project funded by the Ministry of Health and Welfare in South Korea. Seoul National University Bundang Hospital (SNUBH) deployed a HIT system in 2009 that exchanged information - including demographic data, health status, medications, diagnoses, complaints, laboratory results, diagnostic imaging, treatments, care plans, histories and summaries – between SNUBH and referring local clinics and physician practices. This study permitted patients to opt into the study, after which point the hospital physicians would be able to see a special button on the display screen indicating that the patient was a study participant. System logs confirmed that physicians clicked the button and accessed the exchange information at over 99% of first encounters. This indicates a high physician interest in this additional data.

Their investigation compares a group of patients participating in the health information exchange (HIE) with a group of patients who opted out of the pilot program. Clear difference in the charges for various tests and in the length of hospitalization can be observed. Total charges were reduced significantly. These indicators have a clear impact on cost, however the physicians surveyed for this and a previous study by the same researchers indicated that perceived benefits in cost reductions were viewed as far less than the perceived benefits in patient outcomes and quality of care.

The researchers point out that there are several limitations to these findings. First, the cases involved simple referrals between providers and the hospital. Cases in which multiple providers interact simultaneously have yet to be examined. Second, this study did not attempt to measure direct quality of care indicators, which the researchers believe could lead to an underestimation of the benefits of HIT. Third, there were no significant workflow changes as part of the introduction of the HIT. Workflow issues have to be considered, but were not as part of this trial. Some of the benefits could certainly have been related to workflow modifications attendant to HIT adoption. Finally, the researchers warned about limitations on the type of health system studied and the potential generality of the findings. They caution strongly that South Korea has a fee-for-service mechanism, with limited policy initiatives to contain healthcare cost. This certainly impacts the ability to extrapolate results to other locations with centralized health care or significant cost-saving incentives. However, this warning actually implies a strong correlation between the U.S. system and that of South Korea. The system in the United States is a fee-for-service system, and while the debate around health insurance has been vigorous and produced some level of regulation, the discussion of costs for medical services has had far less traction. In this way, it would appear that a study of South Korea's HIT is at least reasonably applicable to HIT in the U.S.

APPLICATIONS OF HIT

Future uses of HIT have been proposed that require additional consideration. One proposed by Fichman, Kohli and Krishnan (2011) is the potential for evidence-based medicine (EBM). EBM promises the ability to make healthcare decisions based on strong, empirical evidence rather than basing decisions on personal habits, tangible and intangible incentives, or medical traditions that could be of various and questionable value. EBM adoption is limited by several factors. One is the lack of information available about the actual efficacy of various treatments. Digital storage of health information could give researchers opportunity to investigate links between treatments and outcomes to arrive at better analyses

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and tools. Another is the time it takes for newly discovered medical information to circulate through the medical community. HIT could facilitate distributing that information to providers or integrating it into automated tools.

Another possibility proposed by Fichman, Kohli and Krishnan (2011) is that of personalized medicine. Personalized medicine uses specific details about a patient's unique physiological makeup to allow earlier and more accurate diagnoses and treatment. Genetic information can be included, identifying risk factors and potential treatments that have been linked to certain genetic traits.

POTENTIAL RISKS AND MITIGATION FACTORS

The most significant risk to HIT implementation, especially implementation of a national registry, is security of private information. According to Chen and Lambright (2016), healthcare information systems (HIS) lag other industries in the development of adequate cyber security, but the information stored in an HIS is of great value to hackers. Their paper states that the healthcare sector reports 340% more information security incidents than other sectors, with 81% of healthcare organizations reporting cyberattacks in the two years between 2014 and 2016. Yet a 2015 industry survey showed only 53% of providers and 66% of insurers report being prepared to defend against a cyberattack.

This perceived risk affects the willingness of patients to participate in this technology. Fichman, Kohli and Krishnan (2011) advise that patients perceive a higher risk of compromise than actually exists, which is quite telling when one considers the actual risk of compromise as explored by Chen and Lambright (2016). Willingness to disclose is critical to this effort, because the effectiveness of treatment and the impact on public policy can be significant. Unless and until security concerns are adequately addressed, any attempt to implement a national health database would be met with skepticism and resistance, undermining the entire effort.

A NOVEL SECURITY MODEL

Chen and Lambright (2016) present an interesting model for potentially securing this sensitive information. Their model presents an autonomic security monitoring system. This system constantly would imbue the registry with a certain level of ability to self protect in much the same way that the human body uses the autonomic nervous system to monitor functions and maintain homeostasis without direct, conscious involvement. Their Autonomic Security Management Framework (ASMF) would contain elements of risk assessment, prevention mechanisms, early warning and detection and limited response actions.

They base this framework on a three-tier architecture whose three tiers are a web tier, an application tier, and a data tier for storing and retrieving information. In this model, the data tier collects data from relevant participants and store the data in repositories. The web tier provides server-based applications for use by providers and patients. The application tier contains mobile web applications that communicate and transfer data and information between the other two tiers. The researchers identified five prevalent attacks that healthcare systems would be susceptible to and identified approaches for each attack. The first attack is scanning attacks, in which hackers would scan devices in the network in order to gather network information that can be used to launch more sophisticated attacks on the system as a whole. The second is spoofing, in which adversaries could impersonate authentic system users to gain access

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

either by masquerading passively as part of the system or masquerading actively as validated system users. The third is injection attack, in which hackers may exploit vulnerabilities in the data tier to insert untrusted data. They could inject malicious information or commands in order to reduce the security of the overall system. Fourth are broken authentication and session management attacks. These attacks exploit vulnerabilities to assume the identity of a valid system user. The fifth type of attack is denial of service or DOS. These attacks make the system unavailable by overwhelming it with bogus requests that overwhelm system resources, blocking legitimate uses.

To counteract these attacks, the authors relied on the use of Risk assessment, intrusion Estimations, intrusion Detection, and intrusion Response (REDR) modules. Risk assessment is divided into five steps. The first step is to characterize HIS, in order to create a "normal region" that represents secure system performance according to relevant metrics like CPU Time, Memory Utilization and I/O Read/Write Rate. Step two is to identify threats and vulnerabilities, from which a protection requirement is generated. From that protection requirement, a comprehensive baseline and system controls are selected and implemented in steps three and four. Step five is to assess and report on the performance of the system. The intrusion estimation module contains system sensors that constantly monitor selected features. Historical values are used to project future values, which are then combined with system models and used to deploy early warnings when measurements fall outside expected normal or predicated parameters. This scan can also execute appropriate mechanisms to respond to attacks as they occur, without waiting for controllers to respond. The intrusion detection module is a data analysis module that analyzes real-time information to detect anomalous system behavior. This must be upgraded frequently to identify attacks sooner and uses live forensic analysis tools to learn unknown attack signatures. New attacks could be automatically detected and added to the catalog of known attacks for future detection. The intrusion response system gives the HIS the ability to self-protect, mitigating impacts of attacks and returning system parameters back to normal. Dynamic response systems allow autonomic management of the system, allowing the HIS to execute lower ranked responses without human intervention.

The researchers were able to design a prototype system with two demilitarized zones that applied the previously discussed REDR model. Users of the system would be able to send requests for information via the Internet, with known malicious requests filtered by the firewall. Only legitimate users and more sophisticated attacks would be capable of violating this barrier. The intrusion estimation model monitors for known and unknown attacks that bypass the firewall, and the intrusion detection model activates prevention mechanisms. Once new attacks are detected, the intrusion response module can autonomously select protection and mitigation mechanisms. A second firewall is added within the ASMF to filter out any internal attacks or illegitimate requests.

This theorized system represents a cost-effective automated method for detecting and responding to low-level, unsophisticated attacks. It is theorized that coupling this method with other complimentary approaches, like honeypots to prevent denial of service attacks, will make HIS more reliable and safe, protecting large volumes of very private and sensitive data. The researchers plan to create an HIS testbed in order to further explore this option.

Chioreanu et al. (2014) propose a hybrid cloud system, with a public and private cloud for storing sensitive information. This particular model proposes using secure communication networks and encrypted certificate pairs to protect sensitive data. Creating a separate administrator for the public cloud and the private cloud, with access restricted to appropriate users, provides additional security. However, this system was modeled on a very small scale – only hosting the documents of a single department within one hospital – and was not tested against malicious attacks. It relies on very simple protection

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

mechanisms, a private network and encryption keys, and the more involved security functions are handled by a system administrator. The use of cloud computing provides greater computing power in a virtual environment which can reduce the reliance on data centers that create a single point of vulnerability.

RELIABILITY OF HIT

Another potential problem that must be addressed is reliability. Gawanmeh et al. (2015) investigated the reliability of current HIS. They examined the different verification models that can be used to identify internal errors that can cause expensive and even dangerous adverse events. This includes not only record-keeping and data management systems, but also wearable technology and remote devices, such as radiation therapy devices and pacemakers. The authors examined studies of HIS that used simulation to verify system performance. Many of these simulations were used to examine performance factors related to scheduling, patient flow, and physician loading as a means of streamlining appointment scheduling. A number of simulations were also used to evaluate security requirements. Authentication, confidentiality, integrity, and privacy are factors that have been tested using simulation. Simulation has also been used to test the performance properties of various medical devices, such as pacemakers and electrocardiographs. Formal methods that use mathematical models to analyze systems are highly recommended for safety critical systems, and when they are automated they can assess complex systems successfully. They have been used for verification of such aspects as efficiency, security, reliability, computer aided diagnosis, medical devices, access policies, and electrocardiogram algorithms. Several research groups have found, in different cases, that developing software using formal methods of verification produces more reliable applications than other methods (Flemming & Huebner, 2013; Newell, 2011; Yusof, Papazafeiropoulou, Paul, & Stergioulas, 2008). Formal methods have also been used to assess trust metrics and survivability. Semiformal methods are based on formal methods but allow for informal semantics. These methods have been used recently in testing HIS, but this is a newer development that has less proof of concept and fewer case studies to confirm their results.

Gawanmeh et al. (2015) concluded, from their literature review, that healthcare systems should not be limited to one particular type of verification, but that multiple methods should be used due to the complexity of these systems and their many components. In particular, they concluded that test case generation would often require informal or semiformal methods due to their lack of source data. Overall systems may be subjected to probabilistic verification because workflow transitions occur with certain probabilities. No standard method for verifying trust could be established, while multiple methods could be used to verify privacy. They encouraged strong data verification, because there is a particularly high risk of missing, incomplete or inaccurate information in HIS. For applications involving large volumes of data, such as cloud-based assessment of complex records like genomic data and medical image processing, formal methods were recommended as potentially enhancing trust.

ISSUES OF IMPLEMENTATION

Barbarito et al. (2012) examined the effects of interoperability standards on regional healthcare in Lombardy, Italy. Over a period of 10 years, the Lombardy Region has set up the Regional Social and Healthcare Information System (SISS, from the Italian acronym) that connects all healthcare providers

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in a region of about 10 million citizens. The approach was twofold – first was to push all hospitals to accept the Health Level 7 (HL7) standard for their individual records and second, to provide a technical infrastructure for interoperability that each individual location could freely and easily access. The Italian healthcare system is a single-payer centralized system in which regional governments cover the expenditures for their citizens. The "Local Healthcare Units" (ASL in the Italian acronym) manage services for a specific geographic area within the region. Each patient refers to a general practitioner (GP) who is chosen from a local unit close to the patient's residence. Pharmacists, GPs, hospital physicians, specialists, nurses, office personnel, and the Regional Call Center personnel use the SISS. At the end of 1970, clinical and administrative data for all living residents were collected on a monthly basis. In 1982 central registries for patients and GPs were created. The unified SISS project was started in 1999. In 2010 the region introduced a lifelong personal health record (PHR) that collected all available documentation of a single citizen, with the goal of making this information available to all providers regardless of the location where a document was produced.

Information flows within single units were standardized and synchronized to the central registry, but by creating middleware that translated standard documents from each location to the central standard. The interoperability specifications were designed in a top-down approach, and any developer must adopt their technology to these specifications in order to implement a system. The system uses a three-level architecture. The Central Registry of Clinical Data does not store individual clinical electronic documents (CED), but rather indexes them according to standard messages sent by the hospitals where the CEDs are generated. A connective infrastructure, the Extranet, is the intermediate level between the central domain (first level) and local practitioners (third level). It is a virtual private network (VPN) with encryption software that provides communication between the different actors. Individual documents are the property of the patient, who can hide specific files or types of files if desired. Individual physicians query the central registry through the Extranet, and then send a request to the front office of the facility holding the specific CED desired, who then returns an electronic copy of the record. A separate process is used for booking appointments. A patient calls the Regional Call Center and a worker at the call center queries the central database to find available appointments, then presents the available appointments to the patient who selects on to confirm. Orders (for treatment, tests, or medications) are communicated from the provider to the facility in similar fashion. If a patient is admitted to a hospital, the change is broadcast to the central repository and visible to all providers. All of the indexed documents are arranged chronologically to create a complete, lifelong PHR, unless a patient indicates that it is to remain hidden. However, hidden documents remain on the local repository.

Workers are identified through personal smart cards, and these cards store user profiles that set the level of access and permissions each user has. The citizen also has a personal health smartcard, which is used for identification and to verify right to access care. It can also be used to book appointments and consult the PHR.

At present, this system functionally connects 10 million citizens and 73,000 healthcare professionals. It manages 4.7 million pharmacological prescriptions and 1.7 million laboratory exams per, 490,000 lab reports, 180,000 radiological reports, 180,000 first aid reports, and 58,000 discharge summaries on a monthly basis. The implementation has been long, and delays were encountered in training developers and testing systems. Healthcare workers required proper training. Many had trouble adapting to carrying their smart-cards with them at all times. However, the decision to use middleware to translate individual records into a standard format avoided pushing a single standard on all of the individual locations and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

making major workflow changes that would have caused problems and frustration among providers and patients. Also, adopting mandatory goals (such as 90% of CEDs digitally signed) for hospital administrators and creating incentive payments went a long way in overcome resistance to change.

Clearly, this study cannot be generalized to the U.S. system without consideration for the different healthcare systems found in each country. The U.S. has a fee-for-service private payer system, while Italy uses a centralized single-payer system. However, this study demonstrates that creating such a system is possible. That is a significant finding – namely that the complexity of healthcare does not rule out the possibility of creating a central and interoperable system for sharing medical records. Particularly significant for the United States is the decision to treat individuals PHR as their property. Once this distinction is made, it is easy to see why providers or treatment centers would be required to make these documents available to any person the individual designates. In fact, the American healthcare system already requires portability of a patient's information freely between providers in the Health Insurance Portability And Accountability (HIPAA) act of 1996 (U.S. Department of Health and Human Services, 2017). The creation of a central interoperability platform is merely a tool to facilitate this already mandated exchange.

Peng, Dey and Lahiri (2014) pointed out that, although there is a great deal of legislative and financial investment in the technology, HIS adoption in the United States lags far behind adoption rates in other developed nations. In the first quarter of 2013, the Healthcare Information and Management Systems Society (HIMSS) reported that one in eight providers still lacked a clinical data repository (CDR), which is the most basic building block of an EHR system, while a large percentage of providers who do have a CDR system have yet to implement other related systems.

Human Barriers to Adoption

Peter Orszag (2008) on behalf of the Congressional Budget Office, published a report on the causes behind this hesitance, which highlights multiple problems. Firstly, the savings from HIT are largely experienced by patients and insurance companies, giving providers little incentive to adopt the technology. Secondly, the possible improvements in quality of care are unlikely to yield any tangible rewards for providers, as they are not likely to be compensated in any way for them. Thirdly, adopting a completely interoperable, national EHR system will increase portability of records, making it easier for patients to move to a competing provider. Fourthly, since so few providers have adopted the technology, there is no pressure on hospitals to follow suit. The lack of standardization adds to the uncertainty and confusion of participating in this new effort. Concerns about implementation, patient privacy, security, loss of control over patient records, and the risk of other unintended consequences further inhibit participation.

Fichman, Kohli and Krishnan (2011) also point out that the healthcare industry is heavily influenced by regulations and competition. In most European countries, a single payer (public) system has the ability to subsidize, exert pressure on, and even directly require providers to share information. It is difficult to induce providers to participate in a PHR system without subsidizing them. Because there is a risk that patients whose records are easily transferred will transfer providers more frequently. Additionally, they mention the hierarchical nature of healthcare. Peng, Dey and Lahiri (2014) explore this in great detail. They examine implementation as a knowledge transfer problem. In their analysis, they find that knowledge transfer in the social network made up of healthcare providers has a great potential to push adoption forward in the U.S. They determined that some hospitals are more inclined to be receptive to knowledge related to HIT and EHR, owing to an existing familiarity and comfort with electronic systems

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

in general. By placing these providers in networks with providers who have yet to adopt HIT systems, they can act as mentors to encourage and guide them in the process, disseminating their knowledge. Groups similar to the Mayo Clinic, the Geisinger Health System, and Kaiser Permanente are models for the sort of provider network that can effectively promote adoption of HIT systems. This effect is stronger and more influential. They are respected as mentors in the medical community. These healthcare networks can spur initial adoption. They are pushing it to the point at which the contagion effect will take over. The contagion effect is that point at which a new technology becomes so common that everyone must adopt it or be left behind. The contagion effect can only be useful once the technology has reached a fair level of adoption. They argue that a continuum exists in which knowledge transfer coupled with incentives, guides early adoption but the contagion effect eventually replaces it as the primary force.

These conclusions appear to be supported by the findings of Adler-Milstein and Jha (2011). They found that for-profit hospitals were less likely to participate in a state HIE than non-profit providers. Hospitals with a dominant market share were more likely to participate than those with a more vulnerable market position. Hospitals that were part of a centralized system, similar to the network groups mentioned in the Peng, Dey and Lahiri study (2014), were also more likely to participate. Those states that had the highest rates of adoption were those that provided financial incentives and regulatory requirements for participation. However, one state with high participation, Maryland, was also credited with creating a central infrastructure that supported adoption.

System Maturity as Barriers to Implementation

Another potential barrier to implementation is uncertainty regarding the maturity of these systems. Peng, Dey and Lahiri (2014) mention that a sense of uncertainty about the technology causes reluctance on the part of providers. Carvalho, Rocha and Abreu (2016) attempt to provide a framework for assessing HIT maturity in their literature review. They reviewed 14 maturity models of HIT developed by various entities including health care companies, research organizations, and academic experts. All fell within one of two general approaches: assessing individual system components in highly specialized models or comprehensively examining hospital information system as a whole. Most did not disclose their design process. Many failed to name explicit characteristics for each stage of maturity. Many used overlapping factors, but gave them different and terminology, and failed to apply weights to these factors. Most models assessed maturity and provided a path of improvement, but not all had a systemic process to move from one level of maturity to the next. The researchers concluded that no sufficient model yet exists with a broad enough scope that covers all areas and subsystems of a healthcare organization, while giving proper weight to varying factors based on relative importance that is supported by rigorous validation methods. In the absence of a maturity model that meets these criteria, it will be difficult to provide proper assurance to new adopters of HIT systems.

DISCUSSION

The matter of adopting HIT, specifically EHRs in the form of a central, interoperable national system, can be addressed as a series of hypotheses that require validation. If these hypothesis can be proven, there is sufficient cause and capacity to pursue a system of this type.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hypothesis 1: Health Information Systems provide benefits that make them valuable to providers and patients.

The list of theorized benefits is attractive – decreased cost and increased reliability, accessibility, and quality of care with fewer adverse events. The case study from South Korea demonstrates that financial benefits definitely exist, in the form of reducing the length of hospital stays and the frequency of medical testing. This study was unable to capture the attendant cost savings from administrative efficiencies, workflow improvements, reduced effects and reduced litigation from those events, but the researchers assert that the pilot system produced cost savings in these areas that increased the overall financial benefits of the system. The study did not explore reliability, accessibility and quality of care. However, surveys of healthcare providers provided a relevant insight. In the pilot program, almost all physicians accessed the data contained in the electronic record when it was made available to them, indicating that this data was useful to and valued by providers. Physicians surveyed also indicated that they believed the quality benefits of the EHR system outweighed the financial benefits, which suggests that the quality effects could be quite significant. Related work by Aron et al. (2011) indicated that automating medical sensing functions reduce the occurrence of procedural and interpretive errors. But this study does not explore the full benefits of an EHR system – only the impacts of automated sensing. Hypothesis one is well supported by the available literature.

Hypothesis 2: Large-scale adoption of an interoperable regional or national Health Information System is possible.

The staggering complexity of the healthcare industry makes the implementation of a single system for managing healthcare information difficult. There are many different players, each interacting with different types of information, and many of the players already have paper systems for record-keeping, or a form of EHR that is unique to their practice or facility. Attempting to dictate a single standard for record-keeping across all facilities would be time-consuming and burdensome. Existing records would need to be translated to the new format and workflows would require significant changes. Also, the amount of data would be immense. Compiling it in a central storage facility would require massive resources at great expense. Such a single data repository would also be a single point of failure. Any compromise or problem at such a repository would cause major problems throughout the system. The case study from the Lombardy region of Italy, however, presents a unique alternative. Their three-tier architecture uses central indexing to locate files stored at their point of origin, and file contents are indexed in a single language which allow information to be recalled in a standardized manner without requiring that files be changed to a single standard format. This reduces the amount of effort to implement the system, removes burdensome workflow changes, and avoids the need for a central repository by leaving the original data in its place. The example given is from a single-payer health system. The underlying concept is flexible and adaptable and its structure is not impacted by financial factors. The only significant modification would be considering the role of insurance contracts in arranging care, as certain insurance carriers only contract with certain providers and that would have to be factored in to any appointments made. Similarly, insurers have individual lists of approved medications that must be considered when writing or filling prescriptions. However, these pieces of information can be treated as fields that require additional indices in the system. The flexibility of the design permits for them to be

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

considered. The current system in Lombardy serves 10 million people. There would still be a need to scale the system in size to meet the needs of America's population of 323 million, the proof of concept system is large enough to be considered relevant. This case study serves as support of this hypothesis.

Hypothesis 3: A central Electronic Health Record system can be successfully secured against electronic attacks.

The work by Chen and Lambright (2016) present a framework that gives the system a level of autonomic protection that can protect against low-level security threats, leaving IS personnel free to dedicate their time and energy to more sophisticated attacks. This theoretical system was designed around a three-tier architecture very similar to what was adopted in the Lombardy region and previously accepted as a support of hypothesis two. However, this design is theoretical and has not yet been tested or verified in the field. The current state of healthcare information security is marked by shortfalls in protections and a large number of successful attacks. However, the proposed autonomic system, coupled with the inherent protections of a distributed storage system as described in the Lombardy region, do have the potential to meet this requirement. This hypothesis remains unproven until the ASMF system designed by Chen and Lambert (2016) can be verified and appropriate validation of the entire system is conducted.

At this point, two of the three hypotheses are supported with valid cases and the final one is considered yet unproven. The third hypothesis could be considered confirmed once adequate validation of the proposed security system is carried out. This makes a discussion of implementation strategy for next steps and additional research opportunities worthwhile.

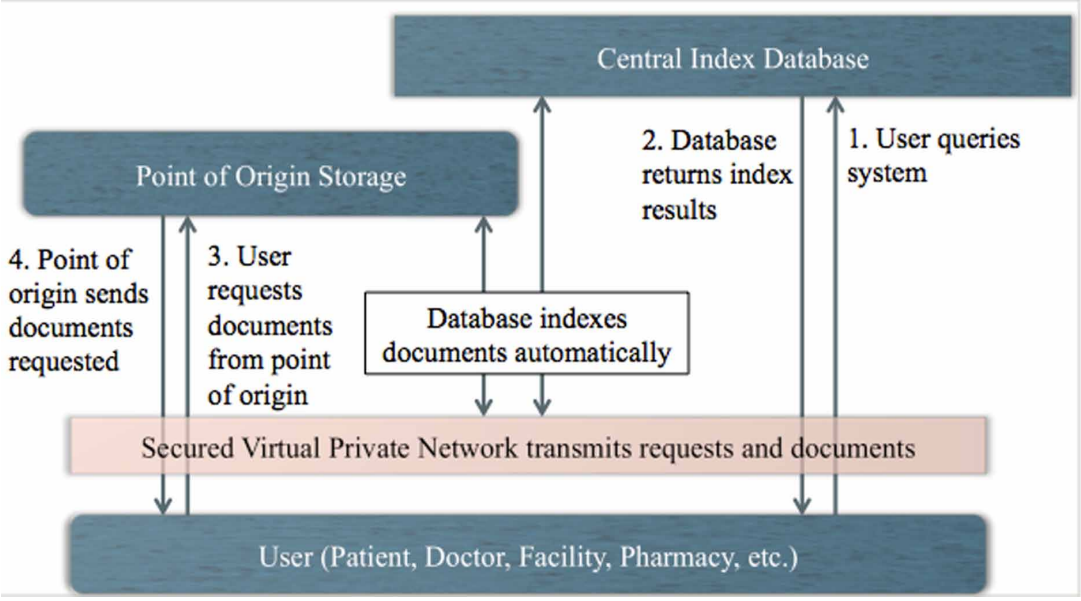
POTENTIAL SYSTEM

We would like to recommend a system similar to that of the Lombardy region, in which data is maintained at its point of origin and accessed through a central index using a query system that permits practitioners to see all entries in a longitudinal patient record. The proposed system differs from the Lombardy system, though. In the proposed system, it has four tiers instead of three. This is because the Lombardy system groups point of origin record sources in with the system users. Although there is overlap between the document sources and the system users, we believe it will be better to separate the two into separate tiers. This will permit additional protections, such as encryption or firewalls, to be enacted around the individual storage sites to further protect them from catastrophic compromise. A diagram of the proposed system is given below. This system would have minimal impacts on workflow, but would require 100% enactment of EHR at all facilities and providers. We have not yet accomplished this goal, but applying knowledge theory through establishing regional or state provider networks to create effective mentoring relationships will be useful. Additionally, as demonstrated by multiple researchers, financial and regulatory incentives will be necessary to overcome the initial financial risks and investments. Once sufficient number of providers have adopted this technology, the contagion effect will capture late adopters, as they are forced to adopt or be left out in the cold.

Sufficient security appears to be attainable, but it will require that the healthcare industry become more serious about their efforts to secure their data against attack. Currently, the industry lags others in giving sufficient attention to this problem. The problem of adoption and the issue of security appear to go hand in hand – as providers have doubts about their ability to secure an electronic system. They are

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Proposed Four-Tier Healthcare Information Exchange System



reluctant to adopt such a system. As fewer providers adopt electronic systems, there is less pressure on the industry to develop effective controls. This is a significant area of further research and development (see Figure 1).

CONCLUSION

As mentioned already, the proposed security system requires verification and validation. Chen and Lambright (2016) articulate plans to validate their model. Additional security verification will be required as the ASMF system develops and is integrated into the overall EHR system. Carvalho et al. (2016) identified various forms of verification and their most appropriate uses. Their work suggests that security testing should use formal, rather than informal or semi-formal, validation methods. Validation of the overall system can be subjected to various methods of validation depending on the individual component of the system to be tested.

No single maturity model exists for assessing the level of development of any potential system. None of the various maturity models that do exist are clear or robust enough, nor do they provide a clear path to developing HIS technology from one level of maturity to the next. A single, clear, actionable maturity model must be developed with clear metrics that are well understood throughout the healthcare industry and among the developers creating HIS. Until this model is developed, it will be difficult to convince potential users to adopt a proposed system.

Another significant gap in the existing body of research is confirmation of the theorized benefits to healthcare outcomes. Hypothesis one was accepted as valid primarily on the weight of financial benefits demonstrated in South Korea. This study could not capture actual benefits to patient health and well-being. Since South Korea's pilot program is a reasonable model for predicting U.S. outcomes, owing to

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the similarity of their health system to our own, it would be particularly useful to return to the SNUBH pilot program with a set of metrics aimed at capturing quality of care. Of particular interest would be any reductions in adverse effects or medical malpractice incidents, as well as other measurable factors like time to diagnosis, time to treatment, treatment efficacy, and missed nursing care. It would also be useful to capture ancillary financial benefits, such as administrative cost reductions.

Pursuing these additional research opportunities would allow us to develop a robust national HIS as the basis for implementing more advanced uses of this technology. Until a national, or at least regional, central EHR can be implemented, the information provided by wearable devices and home health monitoring will not be put to effective use. Similarly, personalized medicine and automated diagnostic tools cannot function in a fragmented system that may not provide complete data for individual patients. Research into developing strong algorithms for these technologies can occur alongside the research necessary to develop EHR interoperability.

REFERENCES

Adler-Milstein, J., & Jha, A. K. (2014). Health information exchange among U.S. hospitals: Who's in, who's out, and why? *Health Care*, *2*(1), 26–32. PMID:26250086

Aron, R., Dutta, S., Janakiraman, R., & Pathak, P. A. (2011). The impact of automation systems on medical errors: Evidence from field research. *Information Systems Research*, *22*(3), 429–446. doi:10.1287/isre.1110.0350

Barbarito, F., Pinciroli, F., Mason, J., Marceglia, S., Mazzola, L., & Bonacina, S. (2012). Implementing standards for the interoperability among healthcare providers in the public regionalized Healthcare Information System of the Lombardy Region. *Journal of Biomedical Informatics*, *45*(4), 736–745. doi:10.1016/j.jbi.2012.01.006 PMID:22285983

Berwick, D. (2011). Launching Accountable Care Organizations — the proposed rule for the Medicare shared savings program. *The New England Journal of Medicine*, *10*(1056). PMID:21452999

Bitton, A., Filer, L. A., & Jha, A. K. (2012). Health information technology in the era of care delivery reform. *Journal of the American Medical Association*, *307*(24), 2593–2594. doi:10.1001/jama.2012.6663 PMID:22735426

Carey, M. A., & Appleby, J. (2013). How do the health law's online exchanges work? What you need to know to get started. Washington Post. Retrieved December 1, 2017 from http://www.washingtonpost.com/national/health-science/how-do-the-health-laws-online-exchanges-work-what-you-need-to-know-to-get-started/2013/09/16/ef7974c8-1988-11e3-a628-7e6dde8f889d_story.html

Carvalho, J. V., Rocha, A., & Abreu, A. (2016). Maturity models of healthcare information systems and technologies: A literature review. *Journal of Medical Systems*, *40*(6), 131–140. doi:10.100710916-016-0486-5 PMID:27083575

Chen, Q., Lambright, J., & Abdelwahed, S. (2016). Towards autonomic security management of healthcare information systems. In *2016 IEEE First Conference on Connected Health: Applications, Systems and Engineering Technologies* (pp. 113-118). 10.1109/CHASE.2016.58

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chioreanu, R., Crişan-Vida, M., Stoicu-Tivadar, L., & Stoicu-Tivadar, V. (2014). Implementing and securing a hybrid cloud for a healthcare information system. In *2014 11th International Symposium on Electronics and Telecommunications (ISETC),* Timisoara, Romania.

Fichman, R. G., Kohli, R., & Krishnan, R. (2011). The role of information systems in healthcare: Current research and future trends. *Information Systems Research*, *22*(3), 419–428. doi:10.1287/isre.1110.0382

Flemming, D., & Huebner, U. (2013). How to improve change of shift handovers and collaborative grounding and what role does the electronic patient record system play? Results of a systematic literature review. *International Journal of Medical Informatics*, *82*(7), 580–592. doi:10.1016/j.ijmedinf.2013.03.004 PMID:23628146

Gawanmeh, A., Al Hamadi, H., Al-Qutayri, M., Chin, S. K., & Saleem, K. (2015). Reliability analysis of healthcare information systems: state of the art and future directions. In *Proceedings of 17th International Conference on E-health Networking, Application & Services*. Boston, MA. 10.1109/HealthCom.2015.7454475

Najaftorkaman, M., Ghapanchi, A. H., Talaei-Khoei, A., & Ray, P. (2015). A taxonomy of antecedents to user adoption of health information systems: A synthesis of thirty years of research. *Journal of the Association for Information Science and Technology*, *66*(3), 576–598. doi:10.1002/asi.23181

Newell, S. (2011). Special section on healthcare information systems. *The Journal of Strategic Information Systems*, *20*(2), 158–160. doi:10.1016/j.jsis.2011.05.002

Nussbaum, A., Tirrell, M., Wechsler, P., & Randall, T. (2010, April 5). Obamacare's cost scalpel. *Business Week*, 64-65.

Orszag, P. R. (2008, July 24). Evidence on the costs and benefits of health information technology. *Congressional Budget Office, Statement before the U.S. House of Representatives*. Retrieved November 2, 2017 from www.cbo.gov/sites/default/files/cbofiles/ftpdocs/95xx/doc9572/07-24-healthit.pdf

Park, H., Lee, S., Hwang, H., Kim, Y., Heo, E., Kim, J., & Ha, K. (2015). Can a health information exchange save healthcare costs? Evidence from a pilot program in South Korea. *International Journal of Medical Informatics*, *84*(9), 658–666. doi:10.1016/j.ijmedinf.2015.05.008 PMID:26048738

Peng, H., Dey, D., & Lahiri, A. (2014). Healthcare IT adoption: An analysis of knowledge transfer in socioeconomic networks. *Journal of Management Information Systems*, *32*(3), 7–34. doi:10.1080/07421222.2014.994672

Rosenbaum, S., & Margulies, R. (2011). Tax-exempt hospitals and the Patient Protection and Affordable Care Act: Implications for public health policy and practice. *National Library of Medicine. Public Health Reports*, *126*(2), 283–286. doi:10.1177/003335491112600220 PMID:21387959

The World Bank. (2017). Health expenditure per capita 2017. Retrieved November 2, 2017 from https://data.worldbank.org/indicator/SH.XPD.PCAP

U.S. Department of Health and Human Services. (2017). Summary of the HIPAA Privacy Rule Retrieved October 25, 2017 from https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Vo, A., & Bhaskar, R. (2012). Health Care Reform Requires Rethinking on the IT Strategy. *Journal of Cases on Information Technology*, *14*(2), 65–72. doi:10.4018/jcit.2012040105

Yusof, M. M., Papazafeiropoulou, A., Paul, R. J., & Stergioulas, L. K. (2008). Investigating evaluation frameworks for health information systems. *International Journal of Medical Informatics*, *77*(6), 377–385. doi:10.1016/j.ijmedinf.2007.08.004 PMID:17904898

This research was previously published in the International Journal of Public Health Management and Ethics (IJPHME), 3(2); edited by Qiang (Shawn) Cheng; pages 39-56; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Section 7
Emerging Trends

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 94
Role of Bioinformatics in Nanotechnology: An Initiation towards Personalized Medicine

Harishchander Anandaram
Sathyabama University, India

ABSTRACT

Recent advancements in bio-computing and nano-technology accelerated the discovery of novel biomarkers in the emerging field of personalized medicine. Personalized medicine deals with disease detection and therapy from the molecular profile of each individual. Personalized medicine is also called as predictive medicine that uses genetic/molecular information to predict disease development, progression, and clinical outcome. In this chapter, we discuss the advantages of using nanotechnology to understand biological systems with an example of the biomarker discovery of cancer. Recent developments in bio computing served as the base for the identification of multiplexed probes in a nano particle. Together we have correlated the bio molecular signatures with clinical outcomes and we have also addressed an emerging field called bio-nano-informatics to suggest an individual therapy for cancer and other diseases.

INTRODUCTION

The nature of self-organization is a key feature of biological system and this feature provides a broader vision for the development of nanotechnology. Molecular diagnosis for identifying human diseases has witnessed an exponential growth in the last two decades and it laid the path for developing advanced technologies for molecular diagnosis and therapy (Ginsburg & McCarthy, 2001). Revolution in molecular diagnosis and therapy was based on discovering novel biomarkers for predicting disease behavior (Little et al., 2008; Jain et al., 2002). Molecular profiling and diagnostics deals with characterizing histological lesions which is heterogeneous in cellular and molecular level of a human disease (Allison et al., 2008). Current technologies like RT-PCR (real time-polymerase chain reaction) and gene microarrays were not

DOI: 10.4018/978-1-7998-1204-3.ch094

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

designed for handling this kind of heterogeneity because they require a constructive preparation of cells and tissue specimen into a homogeneous solution, leading to a loss of valuable information regarding the 3D cellular environment and tissue morphology (Hepper et al., 1984; Liu et al., 2004).The development of nano technology has provided an additional opportunity for integrating morphological and molecular information, which is also followed by observing the correlation between the molecular and cellular changes in disease behavior. Specifically, bi conjugated quantum dots (QDs) have been used to quantify multiple biomarkers in cancer cells and tissue specimens (Steeg, 2008; Wu et al., 2008; Ferrari, 2005; Wang et al., 2008). Nanotechnology can be used for molecular imaging and therapy to improve the efficacy and toxicity profiles of chemotherapeutic agents (Nie et al., 2007; Chan & Nie, 1998; Alivisatos et al., 2004; Michal et al., 2005; Gao et al., 2005). At present, a major challenge in biomedical nanotechnology is to understand the mechanism of interaction between nano particles and biological regimens (blood, cells and organs) under *in vivo* physiological condition. However, the ultimate task is to overcome the inherent limitation to ensure the delivery of nano particles to diseased sites or organs (Gao et al., 2003; Xing et al., 2006; Xing et al., 2007; Yezhelyev et al., 2007; Ghazani et al., 2006). Another challenge is to generate critical studies for linking biomarkers with disease behaviors, such as rate of tumor progression and response of patient to radiation or drug therapy (Sinha et al., 2006; Davis et al., 2008; Jain, 1999; Jain, 2001; Jain, 1998). Finally, we insist on the integration of biomarkers and bio computing with nanotechnology for analyzing a high-throughput data of multiplexed molecular profiles in specimens of cells and tissues.

BIOMARKERS AND THEIR APPLICATIONS

Bio molecular markers can be an altered or mutant gene, RNA, proteins, lipids, carbohydrates and small metabolites, which can be correlated with a biological behavior and its clinical outcome (Wang et al. 2007; Liotta et al. 2000). Most of the biomarkers are discovered on the basis of molecular profiling studies of association or correlation between a molecular signature and disease behavior. Initially, a study was conducted by Gloub *et al.* for classifying tumors on the basis of patterns in gene expression had reported. This study had provided a novel insight into tumor pathology by analyzing the stage and grade of tumor along with the understanding of clinical course and response to treatment (Petricoin et al., 2002; Negm et al., 2002; Ludwig & Weistein, 2005; Golub et al., 1999; Ross et al., 2000). Gene expression studies illustrate the fact that "Molecular signatures of each tumor are inflammatory factors of the original lesion with heterogeneity" (Alizadeh et al., 2000).

The first clinical correlation of gene expression patterns with clinical outcome was reported for the diffusion of B-cell lymphoma (Perou et al., 2000), which is a clinically proven disease with heterogeneity. Whereas, more number of patients had responded well to therapy and survived for a longer period. Hence, it was understood that the variability in a disease progression could be correlated with a specific pattern of a gene expression. Based on this concept, specific molecular portraits were analyzed for the tumor of each individual patient and later, the clinical study was validated by Perou et al. by using an array of clinical samples (Bittner et al., 2000, Dhanasekaran et al., 2001) and it served as the foundation for the discovery of specific expression patterns with the in prostrate, breast, liver and lung cancers (Paik et. al., 2004; Chen et al., 2007; Beer et al., 2002). Biomarkers are divided into predictive, prognostic and therapeutic markers. Prognostic biomarkers predict the natural course of an individual cancer. So, at present it possible to distinguish the indolent and aggressive tumors.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Predictive biomarkers are used to assess the probability of patients benefit in a particular treatment (Hoshida et al., 2008; Hynes & Lane, 2005). For example, breast cancer patients with a normal level of expression of HER2 gene responds well to drug Herceptin, whereas breast cancer patients with up regulated level of HER2 gene responds well to drug tamoxifen (Osborne, 1998; Eisen et al., 1998). Pharmacodynamic biomarkers are used for measuring the short-time treatment effect of a drug in treating tumor (Geraci et al., 2008). In case of clinical applications, biomarkers should be validated by techniques of multiplexed nanotechnology (for proteins) or by RT-PCR (for nucleic acids) (De Souto et al., 2008). Further, we discuss the advantage of using bioinformatics tools for initiating biomarker discovery.

Bioinformatics Tools for Biomarker Research

In earlier days, biomarkers were analyzed on the basis of microarray data using bioinformatics tools with unsupervised learning algorithms (Rehrauer et al., 2007). For example, **Eisen** et al., 1998 had developed a software with several methods of unsupervised clustering and the recent development is a web based application (Rehrauer et al., 2007). Similar methods were used for analyzing the gene expression data from the clinical samples of various patients for identifying the different subtypes of cancer (Pirooznia et al., 2008). Recently, supervised learning methods are used in bioinformatics software for analyzing microarray data in different clinical conditions. For example, MAGMA and ILOOP (Interoven loop) are the two web based applications which is used for analyzing two-channel microarrays (Tarraga et al., 2008; Rainer et al., 2006). ILOOP interface assists the experimental design of a microarray (two-channel) and MAGMA incorporates the normalization procedures in statistics into an application for usability and reproducibility (Gentleman et al., 2004). Most of the web based applications follow a common pipeline of data mining principles. Similarly, GEPAS (Gene Expression Profile Analysis Suite) includes function with steps including data normalization, feature selection and class prediction. In addition to that, CARMA (Comprehensive R and bioconducter-based web) and Gene pattern are recent tools which are being used for microarray analysis (Reich et al. 2006; Michiels et al., 2005; TGO Consortium, 2000; Zeeberg et al., 2003). It was well established that biomarkers are resulted from microarray data analysis on the basis of selection algorithms (Zeeberg et al., 2005; Beissbarth et al., 2004; Carbon et al., 2009; Maere et al., 2005). Apart from sampling algorithms, other databases like Gene Ontology (GO) had given a platform for the evolution of packages like GoMiner, GOStat, AmiGO, BiNGO and GOEAST (Zheng et al., 2008; Frijters et al., 2008; Edgar et al., 2002; Parkinson et al., 2006; Brazma et al., 2001). There was also similar application for searching medline abstracts with specific keywords to display a network structure of GO data (Stokes et al., 2008).

Due to an increase in the accumulation of gene expression data, there is a need for automated packages to mine gene expression data from large repositories like Gene Expression Omnibus (GEO) and Array Express (Ivliev et al. 2008; Zhu et al. 2008; Ochs et al. 2008).There were certain packages constructed for automation but those packages contain certain overlaps in predicting gene expression and hence certain web based applications like Microarray Retriever, GEOmetadb and Array Mining (Hull et al., 2006; Phan et al., 2009; Fogel, 2008). Despite the availability of software, there are challenges in normalization and quality control over gene expression data (Irizarry et al., 2003; Li, 2008; Stokes et al., 2007). Further, there is also a need for translating the list of gene symbols from a feature selection application before the application of a particular GO. Gene Trial Express and Taverna were the workflow application for feature selection (Shei et al., 2006). Another web based bioinformatics resource for biomarker identification was developed by Phan et al. and he named it as omni BioMarker. In this software, biomarkers are

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

identified from several steps which include normalization, feature selection, biological interpretation, validation and clinical prediction. Since there is no single path for biomarker identification, we require unique parameters to address the clinical significance of biomarker.

Clinical Validation of Biomarkers

A study of small set of biomarkers was conducted for the purpose of molecular diagnosis with nanoparticles that are conjugated to antibodies for and targeted therapy (Hernandez et al., 2003; Mora et al., 2005). In particular, multiplexed QD probes can be used to profile a selected panel of biomarker in typical specimen of clinical tissue (Marquez et al., 2005). The use of five to ten protein biomarkers could have a significant impact on disease diagnosis and individualized treatment (Hicks et al., 2001). In this study, four QD-antibody conjugates have been used to recognize and detect four tumor antigens (i.e.) E3 ubiquitin ligase mdm-2, tumor-suppressor p53, zinc-finger transcription factor EGR-1 and cyclin-dependent kinase inhibitor p21/CDN1A. These markers are known to be important in prostate cancer diagnosis and it has been correlated with tumor behavior (Konecny et al., 2003). It is important to note that tumor classification with antigens which are expressed at low levels which can be subjective and therefore it require an experienced observer for identifying a considerable variation. In contrast, quantitative QD measurements allow accurate and user-independent determination of tumor antigens from lowest level of gene expression (Xiao et al. 2004). Thus, the quantitative nature of QD-based molecular profiling could simplify and standardize categorization of antigens of low abundance on intact cells and tissue specimens. This study was proved to be vital for the management of breast cancer. Since the likely benefit of hormonal therapies and trastuzumab is reflected on the quantity of hormone or HER2 receptors (Douglas et al., 2007; Yan et al., 2003; Rothemund et. al., 2004; Chworos et al., 2004).

Integration of Bioinformatics in DNA Nanotechnology and Synthetic Biology

In order to construct an artificial system from biological molecules such as DNA or RNA by the process of self-assembly, there is a requirement to understand the genetic mechanisms behind the action of bio molecules with the help of a high performance computing system. Hence bioinformatics plays a vital role in the integration of synthetic bio systems by the components of nucleic acids (Jaeger & Chworos, 2006; Ponchon & Dardel, 2007; Dittmer et al., 2004; Qian et al., 2005). DNA nanotechnology was established more than 25 years ago by Nadrian Seeman. It is based on the precisely predictable events of molecular recognition between DNA strands with their complementary sequences. In first inception, DNA molecules were used to construct an increasing complex of supra molecular structures ranges from three-dimensional objects (3D) over two-dimensional (2D) lattices (Wilson & Szostak, 1999; Garibotti et al., 2007) to an arbitrary pattern of 2D (Rackham & Chin 2005). From the perspective of synthetic biology, it is particularly interesting to consider the RNA based nanostructures also because of its vital role in the process of transcription. In nature, RNA is rarely used as a structural element. Nevertheless, Jaeger and co-workers could demonstrate a variety of RNA assemblies (Chin, 2006, Seeman, 2005). RNA nano construction process is slightly difficult because of the degradation of RNA. However, recent work suggests that the process of degradation can be avoided by incorporating the structures of tRNA into the RNA transcripts (Bath & Turberfield, 2007; Mao, 1999). In the current era of science, nucleic acid based nano assembly represent the “bottom up” strategy in to capture a part of complexity found in biological self-organization. DNA nano construction is used for finding a DNA sequence which as-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

sembles itself into a structure with a desired thermal equilibrium and annealing process is used to direct the assembly. Some biologists work in contrast to the direct assembly of bio molecules because they believe that most of the complex structures in biological systems are inherently present in the state of non-equilibrium. For an example, cytoskeleton is composed of protein filaments in dynamic equilibrium and the process is constant by the assembly and disassembly of ATP or GTP molecules. In order to explain further, building of complex structures needs the control over the order of assembly in both spatial and temporal arrangements. A biological example for this condition is the complex intracellular rearrangements preceding the cell division or development of multi cellular organisms.

DNA nano devices are proved to be versatile in analyzing machine like assemblies with mechanical properties. Most of these devices are composed of a single or multiple strands of DNA and the molecule can be folded to a designed structure. These devices can change their confirmation by the addition of a strand with respect to the type of nucleic acid and the conformational changes and recently, DNA nano devices has enhanced their versatility by the adaptation of functional nucleic acids (Yurke et al., 2000). In order to achieve the process of self-assembly; the system is subjected to be in non-equilibrium based thermodynamic system with molecular control over the circuits (Yan et al., 2002) and it was stated that the evolutionary process of nucleic acid must be incorporated into the bio based artificial structures (Shin & Pierce, 2004). According to Garibotti et al., the spirit of synthetic biology is the idea of "translation machinery" for nanotechnology. In contrast to the conventional approach by the application of chemicals the artificial translation machinery was developed by Garibotti et al. and mostly they cannot be used to produce an *in vivo* based artificial materials and this lead to the possibility of using the modified ribosomal systems developed by Chin and co-workers (Sherman & Seeman, 2004; Bath et al., 2005; Venkataraman et al., 2007).

Integration of Biological Mechanisms in Computers

Recently, DNA-based molecular recognition has been used to realize various assemblies of machine-likeliness with the property of information-processing and this model is slightly different from nano construction (Adleman, 1994; Ezziane, 2006). These devices are composed of single or multiple strands of DNA, which folds itself into a particular structure and in most of the cases, the structure is designed rationally. The conformation of the devices can be switched between several structures by the addition of DNA or RNA strands or by incorporating a small change in buffer conditions. In this method, the devices can act as simple sensors for the sequences of nucleic acid or the environmental factors which are subjected to the device functioning. The conformational changes can also be utilized to induce the mechanical motion of the device (Yan, 2003; Rothemund et al., 2004; Szostak et al., 2001) and it was also found that the device has the functionality of translational movement (Pohorille & Deamer, 2002; Dittmer & Simmel, 2004; Dittmer et al. 2005; Zhong & Seeman 2006). Finally, the adaptation of functional nucleic acids for DNA nano devices has considerably enhanced their versatility and aptamers were utilized as an autonomous molecular device for computing.

The original work on DNA based computing with an introduction to nanotechnology was initiated by Adleman (Famulok et al., 2001) and certain examples were demonstrated for the processing of DNA based information (Ponchon & Dardel, 2007).DNA model was used for the utilization of certain algorithms to incorporate the process of self-assembly to produce supra molecular patterns (Yingling & Shapiro, 2007; Simmel, 2007) and finally the concept of DNA nanotechnology and DNA computing can be adapted to design novel strategies for the control of transcription and translation. Due to their

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

compatibility in programmable structures, RNA-based devices and control circuits should also be of considerable as the interest for developing the components of artificial cells (Mandal & Breaker, 2004; Sudarsan et al., 2006).

Implementation of DNA Nano Devices

Implementation of an artificial nano device follows the bottom up approach of biological nanotechnology because it utilizes the existing machinery of living organisms on the basis of their metabolism. In *case of an* In vivo operation in a synthetic molecular device may represent a faster road to success. In vivo implementation of molecular devices based on nuclic acids (DNA or RNA) can be achieved in two very different ways *i.e.* (i) Chemical modification of nanodevices, followed by packaging and delivery or (ii) Transfection of cells with artificial genes containing the blueprint for the working of those nanodevices. For example, several DNA based nanomechnical devices were operated using certain effector molecules of RNA *in vitro* (Isaacs et al., 2004; Bayer & Smolke, 2005; Hannon, 2002) and the *in vivo* assembly of such structures has not yet been attempted. In order to achieve an *in vivo* assembly, nanodevices should favour the intramolecular folding of single strands. An another example, we consider the "intramers" (Elbashir et al. 2001;). These intramers are intracellular RNA aptamers, which are produced from an expression system which was genetically engineered to transfer the foreign molecules into live cells. Intramers had been demonstrated to be useful for the *in vivo* targeting of the disease-related proteins. A similar approach is also achieveable for another RNA-based nanodevice. In case of the production of an *in vivo* based nanostructure of RNA, the tRNA-technique may be of great use (He and Hannon, 2004). However, it may not be straight forward approach to transfer the DNA-based assay of *in vitro* technology to an RNA-based assay of *in vivo* system. Currently, certain computational tools are being developed to support RNA-based nano construction (Rinaudo et al., 2007). Recently Simmel et.al gave an overview of the potential in vivo applications of DNA-based nanodevices and they have also addressed the challenges involved in the process implementation (Bansal et al., 2010**).**

Implementation of DNA Nanotechnology in Artificial Ribo Regulators

In the final era of pre genomics and the initial era of post genomics, "riboswitches" have been found to play an important role in the regulation of gene expression in bacteria (Szostak et al., 2001). Ribo switches are natural structures of RNA aptamer and it is being incorporated into the transcripts of mRNA. The transcripts may adopt one of the alternative conformations on the basis of the presence or absence of aptamers. Depending on the conformation, transcription may be terminated or translation of mRNA into protein may be inhibited. In this way, riboswitches exert the control over genes during the process of transcription or translation. It has been shown that, cells perform complex tasks such as logical computations using the circuits of a riboswitch (Pohorille & Deamer, 2002). However, switchable structures made from nucleic acids (DNA and RNA) serve as a central topic in the area of DNA nanotechnology and coupling these switches to gene regulation has been attempted in a variety of different ways. In this case, the construction of artificial riboswitches represents a fusion of ideas from DNA nanotechnology and synthetic biology. In fact, certain artificial ribo regulators were recently devised and implemented in prokaryotic (Forster and Church 2006) and eukaryotic (Seeman, 1982) cells. For example, Isaacs et al. had

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

constructed a part of self-complementary based mRNA molecules, which has the tendency to fold back onto them in such a way that the ribosome binding site was blocked and hence the process of translation was inhibited. This strategy for post-transcriptional regulation had been found in natural ribo switches.

Implementation of DNA Nanotechnology in In Vivo Computing

Another class of a regulatory RNA are termed as microRNAs, which serve as major components in the machinery of natural RNA interference (RNAi) (Chen & Seeman, 1991). In case of RNAi, the process of gene silencing process is triggered by the presence of double-stranded RNA molecules, which are cleaved by the enzyme Dicer into short RNA duplexes termed as short interfering RNAs (siRNA). One of the RNA strands contained in the siRNA duplex is bound by the so-called RNA-induced silencing complex, which induces a site-specific degradation of mRNAs containing a complementary sequence. Excitingly, this natural process of RNAi can be used to knock down genes by the synthetic molecules of siRNA (Shih et al. 2004). Shortly, after the discovery of RNAi, it was also found that the endogenous molecules of regulatory RNA exist and those molecules were currently found to be the microRNAs (Goodman et al., 2005). Recently, Benenson and colleagues utilized the RNAi process to perform logical computations *in vivo* (Gates and Hinkley, 2004). To this end, the computation was implemented into genes, whose expression was regulated by natural microRNAs. When a logical combination of several endogenous molecular inputs was present *in vivo*, the expression was only observed in the reporter gene (coding for a fluorescent protein). Hence, the fluorescence of the modified organisms represents their current logical state which was later termed as the "molecular state." Apart from computing, microRNAs are the prime candidates for coupling artificial DNA or RNA-based nano devices for the processing of life in *in vivo*.

Applications of Bioinformatics and Nanotechnology in Understanding Biological Systems

There are numerous applications to study biological systems by the approach of either Bioinformatics or Nanotechnology but the combination of both approaches have a limited number of a few applications which are very effective and they are (i) Positive feedback gene amplifier (Gates et al., 2005) and (ii) Programmable structure (Gogonea et al., 2001).

Positive Feedback Gene Amplifier

In general, a positive feedback gene amplifier uses the mechanism of a positive feedback to regulate gene circuits by inducing a response to generate binary outputs. Positive feedback method has an advantage over other methods because of the non-requirement of an external factor. At present most of the gene amplifiers of positive feedback uses the mechanism of an auto inducer-independent variant of the quorum-sensing regulator (LuxR). In order to test the potency of the gene amplifier with the mechanism of positive feedback, a case study was conducted by amplifying the output of a one-component of a tetracycline sensor and two-component of the aspartate sensor. In both cases, the positive feedback module amplified the response to their respective inducers on the basis of their dynamic range and sensitivity. In order to construct a circuit of positive feedback, there is a requirement for a transcriptional activator that did not interfere with the regulation of a native gene. In addition to that, there should not

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

be an exogenous inducer to suppress the activator molecule. Based on various studies in the past, an engineered design was made by the three constitutively active variants of LuxR were made to test their suitability in designing an amplifier.

Programmable Structure

Incorporation of nucleic acid based nano devices and computers into a natural or a synthetic biological system is a straight forward approach. In the beginning stage biological cells produce an RNA nano device by transcription. In the next stage, the regulatory mechanisms of genes can be used to control the time of production of nanodevices by a naturally occuring RNA (For e.g., microRNA or RNA) that transcribed from artificial control genes can be used to drive them. Finally, the concepts of DNA nanotechnology and DNA computing can be adapted to devise certain novel strategies for the control of transcription and translation among genes.

FUTURE PROSPECTIVE AND FURTHER CHALLENGES

While addressing the future prospective of biomarker research, there exists a variety of research direction which is particularly promising in biomedical applications but it requires an additional effort for success. The initial direction of research is to design and develop nanoparticles with only one or multiple functionalities. For applications in cancer and other medical conditions, relevant nano particles were used in therapy. Nano particles could also be designed to have novel properties and applications. For example, binary nano particles with two functionalities could be developed for molecular imaging and targeted therapy. Bio conjugated QDs, which have both targeting and imaging functions, could be used for targeted tumor imaging and applications related to molecular profiling. Conversely, ternary nano particles that combine three functions could be designed for simultaneous imaging and targeted therapy. Another direction of research will address the panel of optimizing biomarker using bioinformatics and quantitative molecular profiling with the help of nanotechnology. For example, bio conjugated nano particle probes are used for predicting cancer behavior. The key hypotheses that will need to be tested are that a panel of tumor markers will allow more accurate correlations than relying on single tumor markers. The most important research direction is to investigate the distribution nano particle in ADMET system.

CONCLUSION

Bioinformatics and nano technology aims at the analysis and construction of artificial molecular structures and machines from biomolecules by utilizing the principles of self-assembly and self-organization. In fact, Nuclic acid (DNA and RNA) molecules have been successfully witnessed to realize a wide range of supramolecular structures, nanomechanical devices and molecular computers. Incorporating the potential of “simple” molecular self-assembly and many other properties of biological systems would be desirable for technological systems. For example, environmental responsiveness, robustness, fault tolerance, self-healing, self-reproduction, evolvability, growth and differentiation were some of these properties can probably be realized in compartmentalized nonequilibrium systems containing the networks of complex molecular interactions. Recently, there have been a lot of effort for integrating the components

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

developed on the basis of DNA nanotechnology into larger networks. The genetic nature of its building materials i.e. DNA and RNA, makes DNA nanotechnology compatible with genetic processes. Genes may be used to produce RNA structures.

Gene regulation may be used to control nanodevices or to feed molecular computers and in turn the DNA or RNA switches may be used for unconventional processes of gene regulation. Integration of DNA or RNA-based systems is possible both *in vitro* and *in vivo*. *In vitro*, there have been efforts to incorporate the synthetic gene networks within the artificial cell-like compartments but there are also efforts to directly operate the nucleic acid devices in live cells. Here the utilization of ribo regulators and RNAi processes seems to be promising. In either case, stochastic effects are expected to play an important role in systems behavior and there is a demand for the stable operation of artificial biosystems in small compartments.

The research trends surveyed in this area indicate the fact that a part of the current efforts in DNA nanotechnology and DNA computing will be absorbed in the future prospective of synthetic biology. But the relationship between nanotechnology and synthetic biology is mutual. Current trend of DNA nanotechnology offers the components and concepts for synthetic biology. On the other hand, synthetic biology could be the ultimate "bio nanotechnology" in future.

ACKNOWLEDGMENT

The authors thank the management of Sathyabama University for providing the computational facilities to complete this book chapter.

REFERENCES

Bansal, K., Yang, K., Nistala, G. J., Gennis, R. B., & Bhalerao, K. D. (2010). A positive feedback-based gene circuit to increase the production of a membrane protein. *Journal of Biological Engineering*, *4*, 6.

Adleman, L. M. (1994). Molecular computation of solutions to combinatorial problems. *Science*, *266*, 1021–1024.

Alivisatos, A. P. (2004). The use of nanocrystals in biological detection. *Nature Biotechnology*, *22*, 47–52.

Alizadeh, A. A., & ... (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, *403*, 503–511.

Allison, M. (2008). Is personalized medicine finally arriving? *Nature Biotechnology*, *26*, 509–517.

Bath, J., & ... (2005). A free-running DNA motor powered by a nicking enzyme. *Angewandte Chemie International Edition*, *44*, 4358–4361.

Bath, J., & Turberfield, A. J. (2007). DNA nanomachines. *Nature Nanotechnology*, *2*, 275–284.

Bayer, T. S., & Smolke, C. D. (2005). Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nature Biotechnology*, *23*(3), 337–343.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Beer, D. G., & ... (2002). Gene expression profiles predict survival of patients with lung adenocarcinoma. *Nature Medicine*, *8*, 816–824.

Beissbarth, T., & Speed, T. (2004). GOstat: Find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics (Oxford, England)*, *20*, 1464–1465.

Bittner, M., & ... (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature*, *406*, 536–540.

Brazma, A., & ... (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics*, *29*, 365–371.

Carbon, S., & ... (2009). AmiGO: Online access to ontology and annotation data. *Bioinformatics (Oxford, England)*, *25*, 288–289.

Chan, W. C. W., & Nie, S. M. (1998). Quantum dot bioconjugates for ultrasensitive nonisotopic detection. *Science*, *281*, 2016–2018.

Chen, & ... (2007). A five-gene signature and clinical outcome in non-small-cell lung cancer. *The New England Journal of Medicine*, *356*, 11–20.

Chen, J. H., & Seeman, N. C. (1991). Synthesis from DNA of a molecule with the connectivity of a cube. *Nature*, *350*, 631–633.

Chin, J. W. (2006). Modular approaches to expanding the functions of living matter. *Nature Chemical Biology*, *2*, 304–311.

Chworos, A., & ... (2004). Building programmable jigsaw puzzles with RNA. *Science*, *306*, 2068–2072.

Consortium, T. G. O. (2000). Gene Ontology: Tool for the unification of biology. *Nature Genetics*, *25*, 25–29.

Davis, M. E., Chen, Z., & Shin, D. M. (2008). Nanoparticle therapeutics: An emerging treatment modality for cancer. *Nature Reviews. Drug Discovery*, *7*, 771–782.

De Souto, M., & ... (2008). Cluster cancer gene expression data: A comparative study. *BMC Bioinformatics*, *9*, 497.

Dhanasekaran, S. M., & ... (2001). Delineation of prognostic biomarkers in prostate cancer. *Nature*, *412*, 822–826.

Dittmer, W. U., & ... (2004). A DNA-based machine that can cyclically bind and release thrombin. *Angewandte Chemie International Edition*, *43*, 3550–3553.

Dittmer, W. U., & ... (2005). Using gene regulation to program DNA-based molecular devices. *Small*, *1*, 709–712.

Dittmer, W. U., & Simmel, F. C. (2004). Transcriptional control of DNA-based nanomachines. *Nano Letters*, *4*, 689–691.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Douglas, S. M., & ... (2007). DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 6644–6648.

Edgar, R., Domrachev, M., & Lash, A. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, *30*, 207–210.

Eisen, M., & ... (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 14863–14868.

Elbashir, S. M., & ... (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, *411*, 494.

Ezziane, Z. (2006). DNA computing: Applications and challenges. *Nanotechnology*, *17*, R27.

Famulok, M., Blind, M., & Mayer, G. (2001). Intramers as promising new tools in functional proteomics. *Chemistry & Biology*, *8*(10), 931–939.

Ferrari, M. (2005). Cancer nanotechnology: Opportunities and challenges. *Nature Reviews. Cancer*, *5*, 161–171.

Fogel, G. (2008). Computational intelligence approaches for pattern discovery in biological systems. *Briefings in Bioinformatics*, *9*, 307–316.

Forster, A. C., & Church, G. M. (2006). Towards synthesis of a minimal cell. *Molecular Systems Biology*, *2*, 45.

Frijters, R., & ... (2008). CoPub: A literature-based keyword enrichment tool for microarray data analysis. *Nucleic Acids Research*, *36*, W406–W410.

Gao, X. H., & ... (2005). In-vivo molecular and cellular imaging with quantum dots. *Current Opinion in Biotechnology*, *16*, 63–72.

Gao, X. H., & Nie, S. M. (2003). Molecular profiling of single cells and tissue specimens with quantum dots. *Trends in Biotechnology*, *21*, 371–373.

Garibotti, A. V., Liao, S., & Seeman, N. C. (2007). A simple DNA-based translation system. *Nano Letters*, *7*(2), 480–483.

Gates, T., & ... (2005). Computational materials: Multi-scale modeling and simulation of nanostructured materials. *Composites Science and Technology*, *65*, 2416–2434.

Gates, T., & Hinkley, J. (2004). *Computational materials: Modeling and simulation of Nanostructured materials and systems*. Academic Press.

Gentleman, R., & ... (2004). Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, *5*, R80.

Geraci, F., Pellegrini, M., & Renda, M. (2008). AMIC@: All MIcroarray Clusterings @ once. *Nucleic Acids Research*, *36*, W315–W319.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ghazani, A. A., & ... (2006). High throughput quantification of protein expression of cancer antigens in tissue microarray using quantum dot nanocrystals. *Nano Letters*, *6*, 2881–2886.

Ginsburg, G. S., & McCarthy, J. J. (2001). Personalized medicine: Revolutionizing drug discovery and patient care. *Trends in Biotechnology*, *19*, 491–496.

Gogonea, V., Suárez, D., Van der Vaart, A., & Merz, K. M. Jr. (2001). New developments in applying quantum mechanics to proteins. *Current Opinion in Structural Biology*, *11*, 217–223.

Golub, T. R., & ... (1999). Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science*, *286*, 531–537.

Goodman, R. P., & ... (2005). Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science*, *310*, 1661–1665.

Hannon, G. J. (2002). RNA interference. *Nature*, *418*, 244.

He, L., & Hannon, G. J. (2004). Microrнas: Small RNAs with a big role in gene regulation. *Nature Reviews. Genetics*, *5*(7), 522–531.

Hepper, G. H. (1984). Tumor heterogeneity. *Cancer Research*, *44*, 2259–2265.

Hernandez, I., & ... (2003). Prostate-specific expression of p53 (R172L) differentially regulates p21, Bax, and mdm2 to inhibit prostate cancer progression and prolong survival. *Molecular Cancer Research*, *1*, 1036–1047.

Hicks, D. G., & Tubbs, R. R. (2005). Assessment of the HER2 status in breast cancer by fluorescence in situ hybridization: A technical review with interpretive guidelines. *Human Pathology*, *36*, 250–261.

Hoshida, Y., & ... (2008). Gene expression in fixed tissues and outcome in heptocellular carcinoma. *The New England Journal of Medicine*, *359*, 1995–2004.

Hull, D., & ... (2006). Taverna: A tool for building and running workflows of services. *Nucleic Acids Research*, *34*, W729–W732.

Hynes, N. E., & Lane, H. A. (2005). ERBB receptors and cancer: The complexity of targeted inhibitors. *Nature Reviews. Cancer*, *5*, 341–354.

Irizarry, R., & ... (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research*, *31*, e15.

Isaacs, F. J., & ... (2004). Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotechnology*, *22*, 841–847.

Ivliev, A., & ... (2008). Microarray retriever: A web-based tool for searching and large scale retrieval of public microarray data. *Nucleic Acids Research*, *36*, W327–W331.

Jaeger, L., & Chworos, A. (2006). The architectonics of programmable RNA and DNA nanostructures. *Current Opinion in Structural Biology*, 16.

Jain, K. K. (2002). Personalized medicine. *Current Opinion in Molecular Therapeutics*, *4*, 548–558.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jain, R. K. (1998). The next frontier of molecular medicine: Delivery of therapeutics. *Nature Medicine*, *4*, 655–657.

Jain, R. K. (1999). Transport of molecules, particles, and cells in solid tumors. *Annual Review of Biomedical Engineering*, *1*, 241–263.

Jain, R. K. (2001). Delivery of molecular and cellular Medicine to solid tumors. *Avd. Drug Del. Rev.*, *46*, 149–168.

Konecny, G., & ... (2003). Quantitative association between HER-2/neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. *Journal of the National Cancer Institute*, *95*, 142–153.

Li, C. (2008). Automating dChip: Toward reproducible sharing of microarray data analysis. *BMC Bioinformatics*, *9*, 231.

Liotta, L., & Petricoin, E. (2000). Molecular profiling of human cancer. *Nature Reviews. Genetics*, *1*, 48–56.

Little, P. F. R., Williams, R. B. H., & Wilkins, M. R. (2008). Inter-individual variation in expression: A missing link in biomarker biology. *Trends in Biotechnology*, *27*, 5–10.

Liu, A. Y., Roudier, M. P., & True, L. D. (2004). Heterogeneity in primary and metastatic porstate cancer as defined by cell surface CD profile. *American Journal of Pathology*, *165*, 1543–1556.

Ludwig, J. A., & Weistein, J. N. (2005). Biomarkers in cancer staging, prognosis and treatment selection. *Nature Reviews. Cancer*, *5*, 845–856.

Maere, S., Heymans, K., & Kuiper, M. (2005). BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics (Oxford, England)*, *21*, 3448–3449.

Mandal, M., & Breaker, R. R. (2004). Gene regulation by riboswitches. *Nature Reviews. Molecular Cell Biology*, *5*, 451–463.

Mao, C. D., & ... (1999). A nanomechanical device based on the B-Z transition of DNA. *Nature*, *397*, 144–146.

Marquez, D. C., & Pietras, R. J. (2001). Membrane-associated binding sites for estrogen contribute to growth regulation of human breast cancer cells. *Oncogene*, *20*, 5420–5430.

Michalet, X., & ... (2005). Quantum dots for live cells, in vivo imaging, and diagnostics. *Science*, *307*, 538–544.

Michiels, S., Koscielny, S., & Hill, C. (2005). Prediction of cancer outcome with microarrays: A multiple random validation study. *Lancet*, *365*, 488–492.

Mora, G. R., & ... (2005). Regulation of expression of the early growth response gene-1 (EGR-1) in malignant and benign cells of the prostate. *The Prostate*, *63*, 198–207.

Negm, R. S., Verma, M., & Srivastava, S. (2002). The promise of biomarkers in cancer screening and detection. *Trends in Molecular Medicine*, *8*, 288–293.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nie, S. M., Xing, Y., Kim, G. J., & Simons, J. W. (2007). Nanotechnology applications in cancer. *Annual Review of Biomedical Engineering*, *9*, 257–288.

Ochs, M., & Casagrande, J. (2008). Information systems for cancer research. *Cancer Investigation*, *26*, 1060–1067.

Osborne, C. K. (1998). Drug therapy – Tamoxifen in the treatment of breast cancer. *The New England Journal of Medicine*, *339*, 1609–1618.

Paik, & ... (2004). A multigene assay to predict recurrence of tamoxifen-treated, nodenegative breast cancer. *The New England Journal of Medicine*, *351*, 2817–2826.

Parkinson, H., & ... (2006). ArrayExpress-a public database of microarray experiments and gene expression profiles. *Nucleic Acids Research*, *35*, D747–D750.

Perou, C. M., & ... (2000). Molecular portraits of human breast tumours. *Nature*, *406*, 747–752.

Petricoin, E. F., & ... (2002). Clinical proteomics: Translating benchside promise into bedside reality. *Nature Reviews. Drug Discovery*, *1*, 683–695.

Phan, J., et al. (2009) Improving the Efficiency of Biomarker Identification Using Biological Knowledge. *Proceedings of the Pacific Symposium on Biocomputing* (Vol. 14, pp. 427–438).

Pirooznia, M., & ... (2008). ILOOP - a web application for two-channel microarray interwoven loop design. *BMC Genomics*, *9*(Suppl. 2), S11.

Pohorille, A., & Deamer, D. (2002). Artificial cells: Prospects for biotechnology. *Trends in Biotechnology*, *20*(3), 123–128.

Pohorille, A., & Deamer, D. (2002). Artificial cells: Prospects for biotechnology. *Trends in Biotechnology*, *20*(3), 123–128.

Ponchon, L., & Dardel, F. (2007). Recombinant RNA technology: The tRNA scaffold. *Nature Methods*, *4*, 571.

Ponchon, L., & Dardel, F. (2007). Recombinant RNA technology: The tRNA scaffold. *Nature Methods*, *4*, 571.

Qian, H. (2005). Cycle kinetics, steady state thermodynamics and motors-a paradigm for living matter physics. *Journal of Physics Condensed Matter*, *17*(47), S3783–S3794.

Rackham, O., & Chin, J. W. (2005). A network of orthogonal ribosome center dot mRNA pairs. *Nature Chemical Biology*, *1*, 159–166.

Rainer, J., & ... (2006). CARMAweb: Comprehensive R- and bioconductor-based web service for microarray data analysis. *Nucleic Acids Research*, *34*, W498–W503.

Rehrauer, H., Zoller, S., & Schlapbach, R. (2007). MAGMA: Analysis of two-channel microarrays made easy. *Nucleic Acids Research*, *35*, W86–W90.

Reich, M., & ... (2006). GenePattern 2.0. *Nature Genetics*, *38*, 500–501.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Rinaudo, K., & ... (2007). A universal RNAi-based logic evaluator that operates in mammalian cells. *Nature Biotechnology*, *25*(7), 795–801.

Ross, D. T., & ... (2000). Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genetics*, *24*, 227–235.

Rothemund, P. W., Papadakis, N., & Winfree, E. (2004). Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biology*, *2*, 2041–2053.

Rothemund, P. W. K., & ... (2004). Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biology*, *2*, 2041–2053.

Seeman, N. C. (1982). Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, *99*, 237–240.

Seeman, N. C. (2005). From genes to machines: DNA nanomechanical devices. *Trends in Biochemical Sciences*, *30*, 119–125.

Sherman, W. B., & Seeman, N. C. (2004). A precisely controlled DNA biped walking device. *Nano Letters*, *4*, 1203–1207.

Shi, L., & ... (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature Biotechnology*, *24*, 1151–1161.

Shih, W. M., & ... (2004). A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature*, *427*, 618–621.

Shin, J. S., & Pierce, N. A. (2004). A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society*, *126*, 10834–10835.

Simmel, F. C. (2007). Towards biomedical applications for nucleic acid nanodevices. *Nanomedicine (London)*, *2*, 817–830.

Sinha, R., Kim, G. K., Nie, S. M., & Shin, D. M. (2006). Nanotechnology in cancer therapeutics: Bioconjugated nanoparticles for drug delivery. *Molecular Cancer Therapeutics*, *5*, 1909–1917.

Steeg, P. S. (2008). Heterogeneity of drug target expression among metastatic lesions: Lessons from a breast cancer autopsy program. *Clinical Cancer Research*, *14*, 3643–3645.

Stokes, T., & ... (2007). chip artifact CORRECTion (caCORRECT): A Bioinformatics System for Quality Assurance of Genomics and Proteomics Array Data. *Annals of Biomedical Engineering*, *35*, 1068–1080.

Stokes, T., & ... (2008). ArrayWiki: An enabling technology for sharing public microarray data repositories and meta-analysis. *BMC Bioinformatics*, *9*(Suppl. 6), S18.

Sudarsan, N., & ... (2006). Tandem riboswitch architectures exhibit complex gene control functions. *Science*, *314*, 300–304.

Szostak, J. W., & ... (2001). Synthesizing life. *Nature*, *409*, 387–390.

Szostak, J. W., & ... (2001). Synthesizing life. *Nature*, *409*, 387–390.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tarraga, J., & ... (2008). GEPAS, a web-based tool for microarray data analysis and interpretation. *Nucleic Acids Research*, *36*, W308–W314.

Venkataraman, S., Dirks, R. M., Rothemund, P. W., Winfree, E., & Pierce, N. A. (2007). An autonomous polymerization motor powered by DNA hybridization. *Nature Nanotechnology*, *2*(8), 490–494.

Wang, M. D., Simons, J. W., & Nie, S. M. (2007). Biomedical nanotechnology with bioinformatics - The promise and current progress. *Proceedings of the IEEE*, *95(7)*, 1386–1389.

Wang, X., Yang, L., Zhuo, G. Z., & Shin, D. M. (2008). Application of nanotechnology in cancer therapy and imaging. *CA: a Cancer Journal for Clinicians*, *58*, 97–110.

Wilson, D. S., & Szostak, J. W. (1999). In vitro selection of functional nucleic acids. *Annual Review of Biochemistry*, *68*, 611–647.

Wu, J. M., & ... (2008). Heterogeneity of breast cancer metastasis: Comparison of therapeutic target expression and promoter methylation between primary tumors and their multifocal metastases. *Clinical Cancer Research*, *14*, 1938–1946.

Xiao, Y., & ... (2004). Multivariate search for differentially expressed gene combinations. *BMC Bioinformatics*, *5*, 164.

Xing, Y., & ... (2006). Molecular profiling of single cancer cells and clinical tissue specimens with semiconductor quantum dots. *Int. J. Nanomed.*, *1*, 473–481.

Xing, Y., & ... (2007). Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry. *Nature Protocols*, *2*, 1152–1165.

Yan, H., & ... (2002). A robust DNA mechanical device controlled by hybridization topology. *Nature*, *415*, 62–65.

Yan, H., & ... (2003). Parallel molecular computations of pairwise exclusive or (XOR) using DNA "String tile" self-assembly. *Journal of the American Chemical Society*, *125*, 14246–14247.

Yan, H., & ... (2003b). DNA-templated self-assembly of protein arrays and highly conductive nano wires. *Science*, *301*, 1882–1884.

Yezhelyev, M. V., & ... (2007). In situ molecular profiling of breast cancer biomarkers with multicolor quantum dots. *Advanced Materials*, *19*, 3146.

Yingling, Y. G., & Shapiro, B. A. (2007). Computational design of an RNA hexagonal nanoring and an RNA nanotube. *Nano Letters*, *7*, 2328–2334.

Yurke, B., & ... (2000). A DNA-fueled molecular machine made of DNA. *Nature*, *406*, 605–608.

Zeeberg, B., & ... (2003). GoMiner: A Resource for Biological Interpretation of Genomic and Proteomic Data. *Genome Biology*, *4*, R28.

Zeeberg, B., & ... (2005). High-Throughput GoMiner, an 'industrial-strength' integrative Gene Ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). *BMC Bioinformatics*, *6*, 168.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zheng, Q., & Wang, X. J. (2008). GOEAST: A web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Research*, *36*, W358–W363.

Zhong, H., & Seeman, N. C. (2006). RNA used to control a DNA rotary nanomachine. *Nano Letters*, *6*, 2899.

Zhu, Y., & ... (2008). GEOmetadb: Powerful alternative search engine for the Gene Expression Omnibus. [Keller]. *Bioinformatics (Oxford, England)*, *24*, 2798–2800.

ADDITIONAL READING

Arayne, M. S., Sultana, N., & Qureshi, F. (2007). Nanoparticles in delivery of cardiovascular drugs. *Pakistan Journal of Pharmaceutical Sciences*, *20*, 340–348.

Archakov, A. I., & Ivanov, Y. D. (2007). Analytical nanobiotechnology for medicine diagnostics. *Molecular BioSystems*, *3*, 336–342.

Artigas, R. A., & ... (2008). A novel adrenocorticotropin receptor mutation alters its structure and function, causing familial glucocorticoid deficiency. *The Journal of Clinical Endocrinology and Metabolism*, *93*, 3097–3105.

Barría, M. I., & ... (2009). Analysis of natural variants of the hepatitis C virus internal ribosome entry site reveals that primary sequence plays a key role in cap-independent translation. *Nucleic Acids Research*, *37*, 957–971.

Belting, M., & Wittrup, A. (2009). Macromolecular drug delivery: Basic principles and therapeutic applications. *Molecular Biotechnology*, *43*, 89–94.

Carvajal, C. A., & ... (2003). Two homozygous mutations in the 11 beta-hydroxysteroid dehydrogenase type 2 gene in a case of apparent mineralocorticoid excess. *The Journal of Clinical Endocrinology and Metabolism*, *88*, 2501–2507.

Chen, H. W., & ... (2006). Titanium dioxide nanoparticles induce emphysema-like lung injury in mice. *The FASEB Journal*, *20*(13), 2393–2395.

Chothia, C., & ... (1986). The Predicted Structure of Immunoglobulin D1.3 and its comparison with the crystal structure. *Science*, *233*, 755–758.

Chothia, C., & Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *The EMBO Journal*, *5*, 823–826.

Cieplak, M., & Thompson, D. (2008). Coarse-grained molecular dynamics simulations of Nanopatterning with multivalent inks. *The Journal of Chemical Physics*, *128*, 234906.

Daggett, V. (2001). Molecular dynamics simulations of protein unfolding / folding. *Methods in Molecular Biology (Clifton, N.J.)*, *168*, 215–247.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

De La Calle, G., & ... (2009). BIRI: A new approach for automatically discovering and indexing available public bioinformatics resources from the literature. *BMC Bioinformatics*, *10*, 320.

De La Calle, G., García-Remesal, M., & Maojo, V. (2008). A Method for Indexing Biomedical Resources Over the Internet. *Studies in Health Technology and Informatics*, *136*, 163.

Diekmann, S., & Lindhorst, T. (2002). Dendrimers. *Reviews in Molecular Biotechnology*, *90*, 157–158.

Drexler, K. E. (1992). Nanotechnology: The past and the future. *Science*, *255*, 268–269.

Ehrenfeld, N., & ... (2008). Structure-function relationship between the tobamovirus TMV-Cg coat protein and the HR-like response. *The Journal of General Virology*, *89*, 809–817.

Erickson, B. (2009). Nanomaterial data remain scarce. *Chemical Engineering (Albany, N.Y.)*, *87*, 42.

Eyzaguirre, J., & ... (2004). Importance of the "thumb" for the catalytic activity of endoxylanase B (Xyn B) from *Penicillium purpurogenum. The FASEB Journal*, *18*, C146–C147.

Farokhzad, O., & Langer, R. (2006). Nanomedicine: Developing smarter therapeutic and diagnostic modalities. *Advanced Drug Delivery Reviews*, *58*, 1456–1459.

Fermeglia, M., & Pricl, S. (2009). Multiscale molecular modeling in nanostructured material design and process system engineering. *Computers & Chemical Engineering*, *33*, 1701–1710.

Freddolino, P. L., & ... (2006). Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure (London, England)*, *14*, 437–449.

Freitas, R. (2005). What is nanomedicine? *Disease-a-Month*, *51*, 325–341.

Haddish-BERHANE, N., Rickus, J., & Haghighi, K. (2007). The role of multiscale computational approaches for rational design of conventional and nanoparticle oral drug delivery systems. *Internat J Nanomed*, *2*, 315.

Hamburg, M. A. (2012). FDA's Approach to Regulation of Products of Nanotechnology. *Science*, *336*(6079), 299–300.

Hock, S. C., Ying, Y. M., & Wah, C. L. (2011). A Review of the Current Scientific and Regulatory Status of Nanomedicines and the Challenges Ahead. *PDA Journal of Pharmaceutical Science and Technology*, *65*, 177–195.

Inestrosa, N. C., Álvarez, A., Dinamarca, M. C., Pérez-ACLE, T., & Colombres, M. (2005). Acetylcholinesterase-amyloid-beta-peptide interaction: Effect of Congo Red and the role of the Wnt pathway. *Current Alzheimer Research*, *2*, 301–306.

John, R. (2010). In vivo magnetomotive optical molecular imaging using targeted magnetic nanoprobes. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 8085–8090.

Lee, S., Ferrari, M., & Decuzzi, P. (2009). Shaping nano-/micro-particles for enhanced vascular interaction in laminar flows. *Nanotechnology*, *20*, 495101.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lyshenski, S. E., Kruegel, F. A., & Theodorou, E. (2003) Nanotechnology Bioinformatics: Nanotechnology paradigm and its applications. *Proceedings of the Third IEEE conference on Nanotechnology* (Vol. 2, pp. 896-899).

Moghimi, S. M., & ... (2006). Methylation of the phosphate oxygen moiety of phospholipid-methoxy(polyethylene glycol) conjugate prevents PEGylated liposome-mediated complement activation and anaphylatoxin production. *The FASEB Journal*, *20*(14), 2591–2593.

Moghimi, S. M., Hunter, C. H., & Murray, J. C. (2005). Nanomedicine: Current status and future prospects. *The FASEB Journal*, *19*, 311–330.

Pandey, R., & Khuller, G. K. (2006). Oral nanoparticle-based antituberculosis drug delivery to the brain in an experimental model. *The Journal of Antimicrobial Chemotherapy*, *57*(6), 1146–1152.

Roeske, J. C., & ... (2007). Characterization of the Theoretical Radiation Dose Enhancement from Nanoparticles. *Technology in Cancer Research & Treatment*, *6*(5), 395–401.

Ryan, J. J., & ... (2007). Fullerene Nanomaterials Inhibit the Allergic Response. *Journal of Immunology (Baltimore, MD.: 1950)*, *179*(1), 665–672.

Scott, B. R. (2010). Stochastic Threshold Microdose Model for Cell Killing by Insoluble Metallic Nanomaterial Particles. *Dose-Response*, *8*, 428–447.

Scott, B. R. (2012). Are Some Neurons Hypersensitive to Metallic Nanoparticles? *Dose-Response*, *10*, 37–57.

Singh, K. K. (2005). Nanotechnology in Cancer Detection and Treatment. *Technology in Cancer Research & Treatment*, *4*(6), 583.

Tang, L., & ... (2014). Investigating the optimal size of anticancer nanomedicine. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(43), 15344–15349.

Tao, L., & ... (2011). Shape-specific polymeric nanomedicine: Emerging opportunities and challenges. *Experimental Biology and Medicine (Maywood, N.J.)*, *236*, 20–29.

Tarabout, C., & ... (2011). Control of peptide nano tube diameter by chemical modifications of an aromatic residue involved in a single close contact. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 7679–7684.

Torres, J., Danese, S., & Colombel, J. F. (2013). New therapeutic avenues in ulcerative colitis: Thinking out of the box. *Gut*, *62*, 1642–1652.

Ulbrich, W., & Lamprecht, A. (2010). Targeted drug-delivery approaches by nanoparticulate carriers in the therapy of inflammatory diseases. *Journal of the Royal Society, Interface*, *7*, S55–S66.

Zhang, G., & ... (2009). Tissue Specific Cytotoxicity of Colon Cancer Cells Mediated by Nanoparticle-delivered Suicide Gene In vitro and In vivo. *Clinical Cancer Research*, *15*, 201.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS WITH DEFINITIONS

ATP: Adenosine Tri Phosphate.

Biomarkers: Characteristic biological properties that can be detected and measured in various parts of the body.

Bio-Nano-Informatics: A branch of science which deals with the application of nano technology and information technology to understand biology.

CARMA: Comprehensive R and bioconducter-based web portal.

GEO: Gene Expression omnibus.

GO: Gene Onthology.

In vivo: within the living.

QD: Quantum Dots.

RNA: Ribonucleic acid.

RT-PCR: Real Time Polymerase Chain Reaction.

This research was previously published in Integrating Biologically-Inspired Nanotechnology into Medical Practice edited by M. Amin Bhat, Anima Nanda, and B.K. Nayak; pages 293-317; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 95
Overview of Cellular Computing–Basic Principles and Applications

Amit Das
University of Kalyani, India

Rakhi Dasgupta
University of Kalyani, India

Angshuman Bagchi
University of Kalyani, India

ABSTRACT

Computers, due to their raw speed and massive computing power, have been highly used by biologists to expedite life science research whereas several computational algorithms like artificial neural network, genetic algorithm and many similar ones have been inspired by the behaviors of several biological or cellular entities. However till date both these disciplines i.e. life sciences and computer sciences have mostly progressed separately while recent studies are increasingly highlighting the impact of each discipline on the other. The chapter describes several features of biological systems which could be used for further optimizations of computer programs or could be engineered to harness necessary computational capabilities in lieu of traditional silico chip systems. We also highlight underlying challenges and avenues of implementations of cellular computing.

INTRODUCTION

Computers, especially the modern day desktops and workstations, owing to their massive processing power and speed, have been traditionally used for deeper understanding of several biological phenomena. Techniques essential for high through put processing of genetic (gene sequencing), transcriptomic (micro-array), proteomic (mass spectroscopy) and imaging (microscopy and diagnostic techniques like CT-Scan,

DOI: 10.4018/978-1-7998-1204-3.ch095

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

PET, MRI etc.) data, heavily rely on traditional silicon chip based computing systems (Kuznetsov et al., 2013). However at the heart of every computer, lies a processor which ultimately understands Boolean logic and rely on direct or modified form of Von Neumann architecture. This basic simplistic core of a computer also resembles the behavior of an individual cell which responds to certain input parameters (signals) through the generation of an output based on certain sets of logic. This common feature between a computer and a cell gives rise to a relatively new field of study, termed as 'Cellular Computing' (CC) which focuses on harnessing the power and characteristics of biological entities like cells for the purpose of computation rather than using traditional in-silico network analogy based methods for better understanding of complex systems (Sipper, 1999; C Teuscher, 2012). In case of CC, the word 'computation' does not necessarily restrict its meaning to 'adding numbers' only, rather it is used to describe achievable outcomes through direct or engineered (or synthetic) exploitation of bio-molecular / cellular logic sets which have been evolutionarily selected.

The two following example will provide better understanding of the analogy between computer and cellular logic processing system and a couple of the advantages of CC:

1. The bacterial growth scenario:

Logics like IF/ELSE or AND/OR are a common part of almost every kind of algorithm and rarely requires any introduction. Now it is remarkable to highlight that even the miniature scale cellular biological entity i.e. a bacteria is always responding to its surrounding environment based on such logics and that too is related to its primary aim which is bacterial growth and division (Figure 1). For example IF nutrition is present in the surrounding medium, then continue growth ELSE suspend growth and prepare for tough situation. So it is clear that, without going into any underlying biological pathway details, it is possible to find to similarity between a cell (or bacteria) and a computer. To make a little complex scenario lets add one more factor in the bacterial growth scenario i.e. the temperature. IF nutrition is present in surroundings AND temperature is favorable, continue growth ELSE suspend growth (Figure 1).

The underlying molecular details of this bacterial sensing and growth regulation are also an outcome of series of events of such logic and thereby can be specifically categorized as 'Molecular Computing' which although related but for the better understanding of the readers are discussed in details in a separate chapter and thereby carefully kept outside the scope of this chapter.

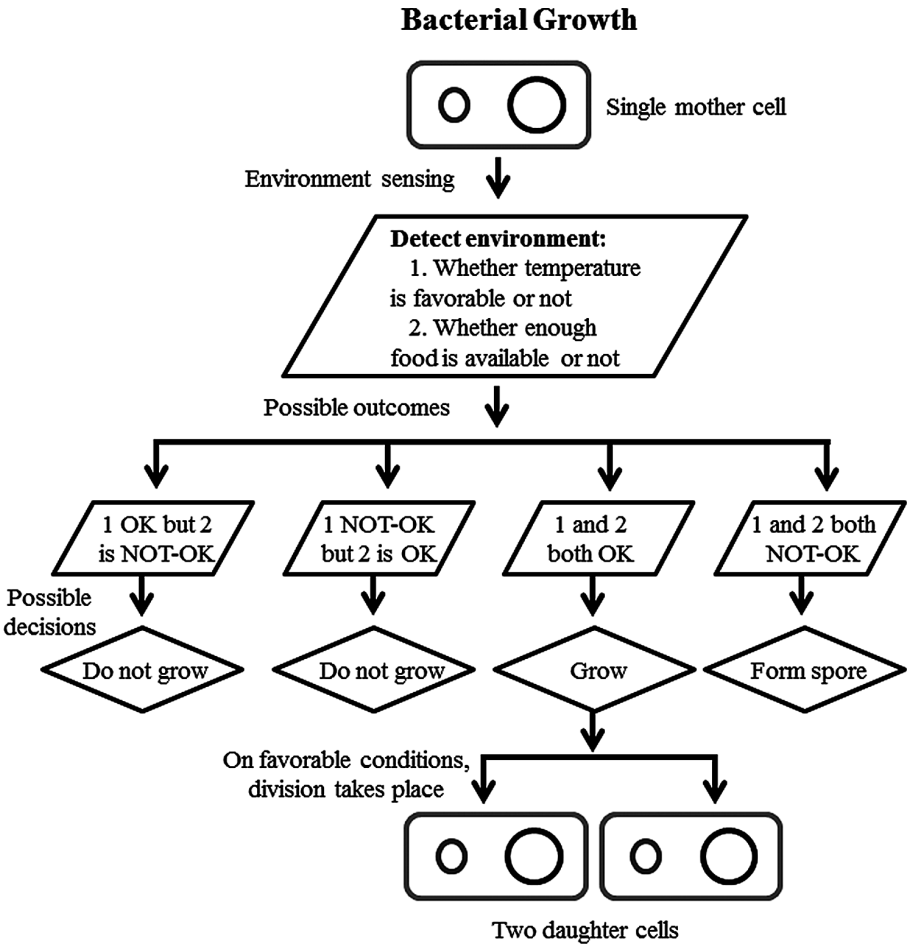
2. The traveling salesman problem:

Before we go in to the details of the problem, let's have a brief understanding of the ways data are processed by a computer. A computer processor primarily works in sequential mode. Although availability of multiple threads combined with massive speed, a computer processor can perform millions of calculations per second but at the heart of this speedy calculator lays a system which can only perform in sequential mode and thereby can only manage sequential inputs. Even in case of present day scenario which show cases the availability of multi-core processor and general purpose graphics processing units (GP-GPU) which are promoted to facilitate miniature to large scale parallelization (depending on the number of processor cores), it must be remembered that they are all essentially arrays of sequentially performing processor which is managed in a smart way (that too by another sets of sequentially performing processor) for a certain extent of performance gain. However separate lines of codes are required

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Flow diagram of the series of logical behaviors which a bacterial cell undertakes to make a decision on the suitability of its division

Like any computer program which assess the input values based on predefined conditions and logics to generate the output, bacterial cells also undergoes similar logical steps to make a decision on their division. First the bacterial sensing machinery senses the environmental conditions (mainly temperature and food) and based on their sensing outcome a decision is made on whether division favorable conditions exist or not. While in reality, several parameters are assessed, a two factor (temperature and food) conditions determination system shows that four different possible outcome of the sensing machinery, only one of which is favorable to continue with bacterial division.



to make a program able to harness this performance benefit and thereby results in its own additional energetic cost. The bottom line of this discussion is that true parallelization of computational architecture is yet to be achieved.

Coming to our point of discussion, the traditional travelling salesman problem is a classic example of efficient management of time and energy and any other humanly associated cost. Now if we look into a bee or ant colony (where lies a social higher archery of queen, king and worker individuals), there will be some individuals, generally termed as foragers (a part of the larger worker community which also harbors individuals like soldiers etc.) who has a primary role of finding food sources for the entire community. Now each community has several such individuals, each of whom, files on its own path of exploration and then returns back and intimidates fellow workers about the place (exact specific direction and loca-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

tion) of available food sources. This can be treated as a classic case of parallelization as discovery of multiple of food sources are quite (and often) possible and each discovery (or exploration) is independent of each other and is not controlled by any central individual along with the fact that multiple discovery results in the increased combined benefit for the entire bee or ant colony. Since the forager has limited travelling ability it automatically has to optimize its exploration path which increases with experience, a process known as learning (similar to machine learning features of a computer). More importantly the worker individuals who will continue to collect food (post discovery by forager who has to travel once) are required to perform multiple trips to make the best of a particular food source. To keep the energy cost minimal it is essential for the workers to travel as minimum distance as possible to reach the food source and return back to nest or hive. Since the forager one was the one who discovered the sources and intimidated its location to the works, the forager is also responsible for optimization of the path to be travelled by the workers. Along with the energy cost the minimal path is also essential to increase the survivability chances as traveling more distance will require more time meaning extended duration of exposure to probable predators (like birds or other such ant or bee hunting animals). Thus from this example it can be concluded that forager bees and their activities are a classic example of parallelization, optimization (in terms of energy and other cost) and learning.

From the above two examples it can be understood that biological entities function in as similar logic driven way (the bacterial growth example) to that of a computer and parallelization, optimization and learning are essential for survival for biological entities (the bee and ant example). Since a computer, which is a logically operating machine aimed at achieving combined greater outcome by performing in parallel with optimal use of energy and time (costs), bears striking similarity to biological systems or entities, the principle of CC focuses in harnessing these in-built capabilities of cellular systems to address day to day challenges either through engineered cells or via optimizations of computer algorithms with inspirations from biology.

In this context it is worth mentioning that there have been huge amount of available information on CC from a computer scientists point to view (Binitha & Sathya, 2012; Fisher & Henzinger, 2007; Navlakha & Bar-Joseph, 2011). Considering the huge amount of bio-inspired algorithms and their diverse field of application, availability of such information it is not surprising. However it should be noted that CC, irrespective of its way of implementation, is based on biological or cellular principles or behavioral properties. Thus it is essential to understand the logical parameters of biological properties or behaviors in order to successfully apply them. Without such knowledge it would be similar to applying technology in the absence of underlying science. In this chapter the authors have provided greater focus to bridge this gap as well as aimed at highlighting CC as an interdisciplinary approach to solve several challenges rather than highlighting CC as a collection of algorithms with inspirations from cellular systems.

BACKGROUND

Robert Hooke, the British naturalist, discovery the basic structure of a cell from dead cork cells using a very basic form of microscope (Hooke, 1665). Since then cells have been explored from almost all possible aspects and with the advent of molecular biology techniques, a clear picture of the underlying cellular machineries were gradually enlightened. However owing to the tremendous complexity of the cell, till today there lays a significant amount of lack of understanding of different cellular behavior even at the level of a single bacterium.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

With the advancement of computer science and hardware, biologists gracefully welcomed the contribution of computers in biological sciences. The contribution of computer in biology is as well diversified as biology itself. Whether it is in the field of microscopy and image processing or in the area of modeling biological systems, with the increase of speed and robustness of digital computers, biologists explored the full potential of these silicon chips for better understanding of the cells, organisms, systems and its interactions with surrounding environment. Efficient use of computers to unravel the secrets of biology gave rise to subjects like Bioinformatics, computational biology, systems biology and so on (Baldi & Brunak, 2001; Kitano, 2001).

On the other hand the in spite of the tremendous complexity of biological systems, cells or multicellular organisms exhibit a very fast response to the changes of its surroundings. While some responses like changes in surrounding temperature or pH, DNA damage are required to be addressed very quickly, other factors like evolution is an ultra-long time scale event which also affects the growth and modifications of biological systems. These characteristics of biological systems like damage control, adaptation, evolution etc. have inspired computer sciences to develop and optimize algorithms for increased stability, robustness and efficient processing of information. Collectively such algorithms became known as bio-inspired algorithms and this gave rise to a field of computer sciences known as biologically inspired or biologically motivated computer sciences (Binitha & Sathya, 2012). The main aim of this steam is not to create an artificial cells or organism but to create efficient data processing programs that shares the time tested characteristics of biological systems.

However the field of CC is not only limited to development of efficient algorithms with inspirations from biology, rather CC also deals with using cells or organisms for the purpose of data processing. The famous McCulloch-Pitts neuron, research on which generated a huge amount of information on neuronal data processing, became a path leader of studies in this area and till today it continues with greater enthusiasm and effort (Arbib, 1995; McCulloch & Pitts, 1943). Later studies by Alan Turing showed that given the availability of sufficient neurons, they can be subjected to logical behaviors as neurons simply behave as NAND gates (Christof Teuscher, 2002; Turing, 1969). Studies by Pattee (1961), Stahl and Goheen (1963) helped in further establishment of the concept proposed earlier by Turing. Sugita argued the fact that since biological systems are merely a well-ordered chemical system, there must be a way of representing them through electronic analogues (Sugita & Fukuda, 1963; Sugita, 1961, 1963). Subsequently the ideas of cellular automata and cells as an information processing system were established. While Schmitt discussed the idea of 'biological macromolecular memory' (Schmitt, 1962), the famous talk 'There's plenty of room at the bottom' by Richard P Feynman pioneered the idea of miniature scale electronics known as molecular electronics or nano-electronics (Feynman, 1960). However at the beginning of CC, his ideas served the purpose of a strong stimulant.

The famous Moore's Law, proposed by the famous Gordon Moore in 1965 (Moore, 1965) has dominated the traditional computational hardware field over the last few decades. Although this law still holds true and is expected to continue its domination to a foreseeable future, the fundamental limitations of the transistor architecture that Moore reported is gradually becoming clear (Kish, 2002) and there now lays an ever increasing need of discovery as well as establishment of an alternative way of information processing. The both wings of CC are expected to fill this void in near future. While bio-inspired algorithms can enable better use of presently available or upcoming digital computer, the information storage, processing and interpretations abilities of cells can reach beyond the physical and fundamental limitations of traditional silicon chips.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DIFFERENCE WITH MOLECULAR COMPUTING

Molecular computing (MC), similar to CC, is also another field of interdisciplinary studies which deals with the process of using biological macromolecules like proteins or DNA for the purpose of computing i.e. storage, analysis and processing of information (Benenson, 2012; Fu, 2007; Lai, Sun, & Chuang, 2014). However there lies some fundamental difference between MC and CC. CC has a two way approach which has two primary goals – i) use of certain characteristics of biological cells or systems for the purpose of optimization of computer algorithms and ii) engineering of biological building blocks i.e. cells for the purpose of computation. On the other hand MC is a one way approach with the primary aim of engineering biological macromolecules for computing purposes.

Since a cell is composed of hundreds of thousands of macromolecules and both CC and MC aims at engineering them for computational purposes, MC apparently seems to be a part of CC. However that would be an incorrect assumption considering their underlying basic principle. In case of CC, the entire cell is treated as a unit while its underlying mechanisms are largely ignored whereas in case of MC, individual protein or DNA molecules are taken in to consideration. In CC, the behavioral or physiological changes of the entire cell, in response to certain input signal (mostly from the surrounding environment) is considered whereas MC takes into account changes in molecular level, sometimes extending up to individual bases of DNA.

Beyond this difference of order of magnitude (molecular level) between CC and MC, CC has another wing where several computation algorithms take inspiration from the behavior of biological cells for development or optimization purposes (a separate section discussing this aspect of CC is discussed later in to this chapter). However extension of macromolecular behavior, especially changes beyond their physic-chemical properties, for optimization of computational algorithms have not been well documented yet.

INHERENT CHARACTERISTICS OF CELLULAR SYSTEMS

As reported by Michael Conrad, a pioneering figure in the field of bio-inspired and cellular computing, cellular organisms have a different way of information processing compared to the digital computers. Traditional silicon chips based digital computers accomplish this task based on several layer of abstractions whereas living entities make the best used of physic-chemical properties of their constituent materials (Conrad, 1989). Cellular computing tries to bridge this issue of the way of information processing through optimization of natural and artificial systems by drawing inspirations from real life biological entities like a cell or an organism.

Thus for proper understanding of the impact of cellular systems on cellular computing, it is essential to understand several cellular properties that have been or can be further exploited for the progress of cellular computing.

Uniformity

Now a day, it is well known that every human, irrespective of their origin, are 99.9% identical. Not only humans, every species that exists on earth share such high degree of similarity of its genetic content among the members of that particular species. This kind of uniformity is essential for maintaining species identity as such compatible genetic content ensures the feasibility of reproduction only across the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

members of a species and thereby eliminates inter-species reproduction. The point to be noted from this kind of specific uniformity of biological systems is the level of security employed with passage of flow of information. Since production of offspring from inter-species courtship is rare and even in those rare cases is mostly infertile, flow of genetic information reaches an automatic terminus ensuring zero catastrophic effect. If such level of uniformity, paired with inbuilt security, can be employed in to computer systems, there can be massive impact on information exchange protocols.

Along with the security benefit of having species specific uniformity, there are several other benefits of having such uniform internal contents. The uniform nature of evolutionarily conserved biochemical pathways makes it easier to predict the nature of changes that can result from alterations of surroundings. Like a shift to non-permissible temperature, all bacteria, existing in that particular growth medium) will seize their growth immediately and this will result in an immediate halt on utilization of food sources which would have otherwise gradually decreased to make the bacterial growth sustainable. Thus the uniform behavior of bacteria facilitates easy modulations of their growth just by alterations of only one parameter (temperature in this case). Uniformity of any system makes it easier to control. In order to maximize proper utilization of computational and its associated resources (like power consumption), uniform behavior of computer programs are essential. However such features, till date, are far from implementation and there rarely exists any resource utilization uniformity across softwares or hardwares.

Scalability

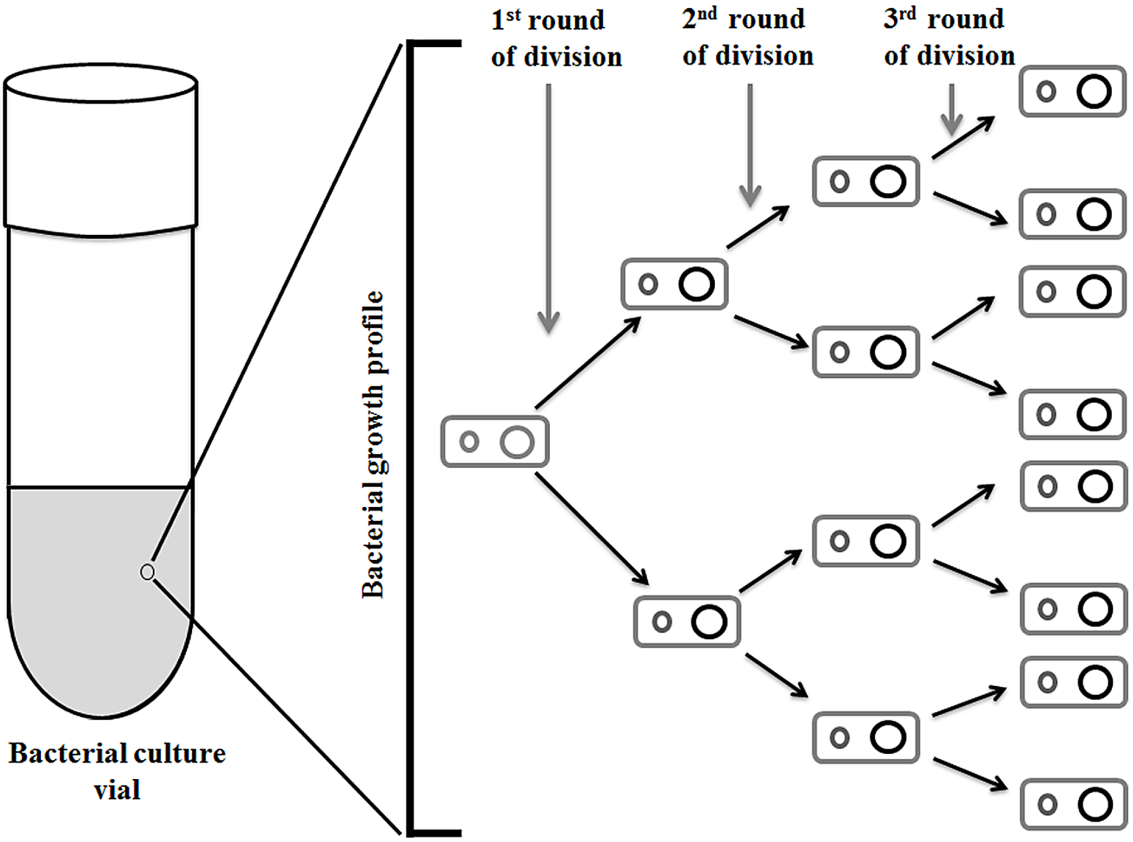
Scalability is an integral part of every cell and it exists in two forms i.e. ability to maximize the available resources as well as its ability to accommodate any increased task in response to changes in the internal and external conditions. When environmental conditions are favorable (as mentioned in Figure 1), a bacterium (for example *Escherichia coli*) can generally divide in 20-30 minutes. With this rate of growth (average of 2 divisions per hour), an overnight grown culture (standard 12 hours growth period which is sufficient enough for the growth curve to just reach a stable state) can contain slightly more than 16 million bacteria (precisely 16777216) even when originating from a single bacterium (Figure 2). Keeping aside the synthesized protein, just to accommodate the genetic material of these 16 million bacteria, around 70 million based of nucleotides have been synthesized and used up for replication (without any error). Thus it can be clearly envisioned, without much effort that such rate of scalability to maximize the available resources is extremely hard to replicate with any existing computational systems but with its implementation, even a fraction of this level, massive increase in resource utilization by computers can be achieved.

Parallelization

Biological systems are massively parallel and parallelism exists at all cellular levels i.e. within a cell, among cells and across multi-cellular systems. Several molecular pathways keep on undertaking their own chains of reactions independent to each other. For example glycolysis and TCA cycle (which are the main energy producing pathways) can simultaneously take place along with DNA replication or transcription. Similarly two cells of an organism can perform different tasks like nephrons facilitate filtering of excretory fluid while rod or cone cells empower the ability of vision despite the fact that both types of cells exist in the same body and share same genetic material as well as similar bio-chemical pathways. As mentioned in the beginning, the forager bees function in parallel and communicate with

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. The growth of a bacterial cell
Bacterial cells (like E. coli) undergo binary fission where at each division the mother cell gives rise to two daughter cells. When favorable conditions are available, such binary divisions can generate 16 million cells from a single mother cell over a period of 12 hours. The chain of such bacterial division is shown up to first 3 divisions which can generate 8 daughter cells.



separate sets of worker bees even though they live in the same have. Thus it can be clearly envisioned that parallelism which is an integral part of any cellular or biological systems, exists at all levels and evolutionary selection have perfected it from all possible aspects. Despite relying on similar bio-chemical pathways, the parallelism exerted by cellular systems is very hard to replicate by traditional computers which mainly rely on von Neumann architecture. However parallelization is such a characteristic of biological systems which computer scientists and engineers should look in to with greater enthusiasm as efficient use of this property both at software and hardware level can increase computation speed and capabilities several thousand folds.

Robustness

Cells have tremendous capability to handle perturbations of its surroundings i.e. cells exhibit significant level of robustness both internally and externally. In terms of internal robustness cells employ several mechanisms to ensure the sequential integrity of its DNA (Papamichos-Chronakis & Peterson, 2013; Sirbu & Cortez, 2013) as any change, some of which even at a single nucleotide level, can have significant effect on the overall system, for example mutations in the p53 gene (Girardini, Walerych, & Del Sal, 2014;

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Muller & Vousden, 2014; Powell, Piwnica-Worms, & Piwnica-Worms, 2014; Yeudall, 2014) or sickle cell anemia (Pace, Ofori-Acquah, & Peterson, 2012). Not only is the maintenance of integrity, compensation and fault tolerance are another important part of cellular robustness. Neurons have developed specific mechanisms to ensure only qualified (above certain threshold) external stimulus are eligible for neuronal transfer along with ensuring no loss of signal intensity as well as proper representation of each and every signal variations with accurate reproduction of type and intensity. Present day computers feature extreme information processing capabilities. However the speed and accuracy with which biological systems filter out noises and extract the required information still remains unmatched and computer science will definitely be benefited if such level of robustness (along with speed) can be employed.

Networking

Networking is an important aspect of both local and global information exchange system, especially in present days. Proper understanding of the network structure and dynamics helps in efficient management of the overall network. Like the computer network, biological cells also form a network where they can communicate with each other. Not only local cell to cell communications, where the communicating cells are connected to each other, long distance cellular communication is facilitated by biological messenger systems like the hormones. Microscopic organisms like slime molds also exhibit efficient network management systems like the slime mold food tunneling network (Latty & Beekman, 2011; Yip, Beekman, & Latty, 2014) where each available food source is connected to each other either via direct connection or through intermediated junctions. Evolutionary forces have made selection to favor only those network connections that ensure minimal cost and high robustness. Understanding the nature of such selections would have in development of efficient computer networks that in turn would ensure maximum utilization of networking components and fast and secure data exchange over such networks (Marwan, 2010).

Discreteness

Despite behaving very synchronously, cells at its core are discrete in nature where each individual cell relies on its own machinery. Such discreteness enables each cell to survive, function and reproduce independent to each other. Even in a multicellular system, where each and every cell must communicate with each other and function in an orchestrated manner, each cell maintains its discreteness. Such discreteness are essential for the survival of the cell in an environment where competition for resources (like food) are always at its limit and this feature in turn enables selection to select the best performing unit. Although each and every computer node of a network or a single computer is a discrete system, it, on its own, cannot undertake ways to survive in a resource limited environment. Intelligent algorithms can enable computers to maintain their discrete identities even when competing for resources with each other.

Speed and Specificity

Cells exhibit remarkable speed and precision with almost all of its functions. While the speed of neurons which is around 500 km/h, can be said to be significantly slower than the speed of a computer which in theory operates at a speed of electricity / light, when the complexity of biological information is taken it amount, the speed of biological phenomena is found to have no match at all. Even with present day super computer architecture, simulations of a single biological cellular activity is still out of question, whereas

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

cells keep on doing their work without any fail. The entry of a virus in to a cell which in general takes a few seconds can keep present day supercomputer busy for a handful of days trying to simulate this event even under much simplistic environment conditions. On the other hand each and every DNA sequence is composed of only four different nucleotides (adenine, guanine, cytosine and thymine), transcription factors are able to bind to their specific target sequences and that too can be found across millions of base-pairs of DNA of a cell, and even a variation of single base can be identified by the transcription factors and the time scale of such events are in to micro-seconds order. Another point to be noted in this context that cells only employ a fraction of its resources to perform such jobs and several such jobs continue to run in parallel within a cell at all times. When such factors (number of base pairs to search, identification of exact match, time taken to perform such job, amount of resource available for the job), it can be clearly understood that the combined speed and specificity of a cell still remains unmatched and there is a long way to go before achieving such level of combinations of precision and speed with traditional digital computers.

In-built encryption

Data encryption and flow of information in encrypted format across the World Wide Web is a constant field of improvement in present day. With the ever increasing threat of mishandling of personal and public data and storage of such data over cloud based storage system, encryption has received its highest level of attention over the last few years. But if we look into the cellular system, it is fascinating to find that encryption has always been an integral part of biological information storage as well as the passage of information across biological networks. Taking into the consideration of the central dogma of molecular biology (i.e. DNA > RNA > Protein), it should be noted that proteins are the primary structural and functional unit of any cell. However the protein coding information is stored into the DNA which has a completely different building block compared to the proteins (nucleotide bases for DNA while amino acids are the primary constituents of protein). Also the passage of information from DNA to Protein takes place via an intermediate stage of RNA. Also there are two other features of this biological information encryption system – i) requirement of only a handful of macromolecules and ii) one directional flow of information. For the first point it should be noted that DNA / RNA are made of only four different types of nucleotides (adenine, guanine, cytosine and thymine for DNA and uracil replacing thymine for RNA) while proteins are made of only twenty different types of amino acids. Thus all the genetic information of each and every species that presently exists on earth or ever existed can be made up of only four different nucleotide bases. Till date no such computer storage system exists that can hold such enormous data with only four different types of constituents and that too in an encrypted format. Same argument also exists for the proteins which are primarily formed by different combinations of 20 building blocks. To further understand the level of encryption that exists in biological systems, it should be noted that protein sequence can be interpreted from the DNA sequences whereas the reverse is almost (and practically) impossible to deduce due the redundancy of genetic code. Thus the decryption can only take place in one direction which is also favored by evolution as no biological machinery exists to undertake the reversed flow of information from proteins to RNA / DNA. If such level of encryption can be achieved by computer systems that can store information with such handful of elements, there would be a massive jump in terms of space required to store information as well as in terms of level of security of information storage.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Optimized Redundancy

Availability of redundant data sources ensures un-interrupted flow of information in case of failure of any node. However to determine the level of redundancy in data storage is a very critical parameter. Cellular systems, over the course of evolution, have developed a highly optimized redundant level of information storage. It is often seen that cells keep multiple copy of genes affecting exact or similar functions. This ensures that in case of deletion of any section of a chromosome or gene, it is functionally complemented by up regulating the expression of these duplicated genes (Gu et al., 2003). However not all genes are maintained in multiple copies. Only those genes which are highly essential for the basic survival of the cell are kept functionally redundant. Thus the challenges for network system specialists or data scientists is to determine which local node or which section of data is to be maintained redundantly to ensure combined benefit of the entire network. Detailed understanding of biological systems, especially gene disruption and duplication machineries and their underlying evolutionary forces would definitely help computer science in this aspect.

Uniqueness Identifying Ability

Uses of bio-metric security features like finger print, facial passwords have become a common feature now a day. While such software tools have become easily available only over the last decade, they existed across biological systems since its origin. Also biological systems execute this unique identifying ability with far less training than that is required by any computer program which often require hundreds or thousands of examples as training sets (Thorpe, Fize, & Marlot, 1996; Zhu, Rogers, Qian, & Kalish, 2007). Thus there lies an opportunity to employ the discriminating properties shown by biological systems to develop an algorithm which is capable of learning in one go (Fe-Fei, Fergus, & Perona, 2003; Fei-Fei, Fergus, & Perona, 2006).

ALGORITHMS INFLUENCED BY BIOLOGICAL ENTITIES

As previously mentioned, several computational algorithms are inspired by biological entities. For the ease of understanding, this section can be divided in to two sections – i) Algorithms inspired by cellular properties / organism behavior and ii) Cellular behavior or biological phenomenon with impact on algorithm optimization. While these two sections apparently seems to be similar, the difference lies in the ways of implementations of the interplay between biology and computer science – i) one algorithm may take inspiration from more than one biological principle while ii) one single biological phenomena may have several characteristics which have been utilized by more than one algorithms.

Bio-inspired algorithms can be broadly classified in to three groups – i) evolution influenced, ii) inspired by ecological systems and iii) swarm based (Binitha & Sathya, 2012).

Among these three groups evolutionary algorithms (EA) is the most well established and is based on evolution of biological organisms. EAs are non-deterministic in nature. The EA family consists of genetic algorithms, genetic programming, evolutionary strategy, differential evolution and paddy field algorithms (Back, 1996; Binitha & Sathya, 2012).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The natural ecosystem comprises of several biotic and abiotic factors which interact with each other in numerous ways. There also exist several levels of competition i.e. inter-species and intra-species completion for food and other non-renewable resources. Along with competition, there lies other modes of interaction like cooperation, predation etc. The natural eco-system thereby represents an ideal example of different level of interactions across a wide range of entities. These complex interaction networks and the flow of energy across the ecosystem have inspired several algorithms like the bio-geography based optimization algorithm which information sharing across candidate solutions (Simon, 2008). The idea of co-evolution of symbiotic organisms has been utilized in the optimization of PS2O algorithms (Chen & Zhu, 2008). Numerical stochastic search algorithms like the invasive weed optimization (IWO) is inspired by the ecological distribute of weeds and their colonizing properties (Mehrabian & Lucas, 2006).

While EAs constitute the most established group of bio-inspired algorithms and eco-system influenced ones takes in to account several properties of complex interactions of biological systems, the remaining type of the bio-inspired algorithms i.e. swarm based algorithms or optimization techniques exhibit maximum diversity. Swarm based algorithms or swam intelligence (SI) is relatively new (Bonabeau, Dorigo, & Theraulaz, 1999). While EAs underlying principle is the genetic adaption of organisms, collective social behavior of a group of organisms form the base of SI. The five fundamental principle of SI are – proximity, stability, adaptability, diverse response and quality (Kennedy & Eberhart, 1995). Intelligent water drops (IWD) algorithm is based on the observation of behaviors of water drops in a natural river (Shah-Hosseini, 2007, 2008). The behavior of human immune system in response to foreign antigens exhibit powerful learning and memorizing abilities, paired with high level of robustness and feature extraction capabilities. These properties have been exploited in to the artificial immune system (AIS) algorithm which is based on the clonal selection principle of the immune system (Dasgupta, 1999). Other than IWD and AIS, the foraging (mainly for food) activity of birds has inspired the development of particle swarm optimization (PSO) algorithm. The social nature of ant and bee colonies has given rise to algorithms like ant colony optimization (ACO) and artificial bee colony (ABC) algorithms (Blum, 2005; M Dorigo, Maniezzo, & Colorni, 1996; Marco Dorigo & Stützle, 2004, 2010). The schooling behavior of fishes has a tendency to sacrifice local minimum to active global success. This natural schooling behavior have inspired the development of fish swarm algorithm (FSA) and primarily exhibit three properties – search for desired resource (like food in nature), exhibit swarming to counteract any possible threat and follow to achieve a successful outcome (Azizi, 2014; X. Li & Shao, 2002).

Since several other chapters of this book incorporate a detailed discussion across almost all the above mentioned bio-inspired algorithms, they have been precisely limited to this point. Table 1 provides a brief description on the most well established major bio-inspired algorithms while Table 2 provides further insight in to some other biological phenomena that have been utilized by computer scientists for algorithm development and optimization purposes but have not been well categorized yet.

CC AND SYSTEMS BIOLOGY

At the very base of CC, its foundation is laid on understanding of the cellular system, especially its underlying logical behaviors. Unlike molecular computing, CC takes in to account the entire cellular system and in this aspect system biology plays an important role. With reductionist approach, understanding individual biochemical pathway or the functions of its individual components are of prime focus. However it should be noted that cells undertake each and every pathway in a very logical manner

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. List of major algorithms inspired by different behaviors and characteristics of cellular or biological entities

Name of Algorithm	Underlying Biological Principle	Applications	Basic Characteristics	References
Artificial Neural Network	Structure of biological neuronal network and learning and processing ability of neurons.	Prediction or forecasting applications, Pattern recognition,	Massively parallel, distributed architecture, learning and generalization ability, fault tolerance	(Jain, Mao, & Mohiuddin, 1996) (Bishop, 1995)
Genetic Algorithm	Natural selection, evolution (select only the best, everything else is discarded)	Automated design, Biological sequence manipulation, evolutionary game theory, robot trajectory, strategy planning	Global and blind search methods, probabilistic transition, parallel implementation	(Goldberg, 1989)
Ant Colony / Travelling Salesman Problems	Figuring out the least path of travel	Task scheduling and assignment, vehicle root optimization, edge detection and linking, network routing	Global, parallel, learning ability	(Marco Dorigo & Stützle, 2004)

Table 2. List of cellular or organismic properties/behaviors that have inspired optimization of several computational algorithms

Cellular Property / Organism Behavior	Inspired Algorithm	Advantage(s)	Disadvantages(s)	References
Fly Sensory Organ Precursor Cells	Finding out maximal independent sets	Distributed, networked, stochastic, modular, robust	Non-adaptive	(Afek et al., 2011; Alon, Babai, & Itai, 1986; Luby, 1985; Rendel, Sheldon, & Finlay, 1965)
Regulation of Gene Expression	Optimized redundancy, fault tolerance	Distributed, networked, adaptive, modular, robust	Non-random	(Balaji, Iyer, Aravind, & Babu, 2006; Gitter et al., 2009; Gu et al., 2003; Guerraoui & Schiper, 1996)
Echolocation of Bat	Trajectory optimization, locating / positioning of movable objects	Robust, adaptive	Non-random, localized	(Yovel, Falk, Moss, & Ulanovsky, 2010)
Foraging of Slime mold	Network path, connection and efficiency optimization	Distributed, networked, stochastic, adaptive, robust	Less-flexibility	(Marwan, 2010; Tero et al., 2010; Watanabe, Tero, Takamatsu, & Nakagaki, 2011)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

with an aim of achieving greater combined benefit for the entire cell. Thus the reductionist approach has serious limitations because deeper knowledge about a specific pathway does not necessarily, which is often the case, reflect the overall outcome which is actually an outcome of several inter-connected (some serial and some parallel) pathways. Hence understanding the entire cellular system is of prime importance in order to take inspirations to apply CC for the betterment of daily life challenges. In this context the systemic approach takes in to account the limitations of the reductionist approach by looking in to interaction networks and the biological system properties which are an outcome of such interactions.

Gene regulatory network are the very basic but essential form of interaction networks of any biological system (Davidson & Levin, 2005; Lee et al., 2002). Qualitative models like random and probabilistic Boolean network, network motifs and asynchronous automata are associated with the modeling of gene regulatory network (Cardelli, 2007; Hecker, Lambeck, Toepfer, van Someren, & Guthke, 2009; Istrail, De-Leon, & Davidson, 2007). In another way the entire gene regulatory network can be represented as a 'genomic computer' which can potentially enlighten computation that takes place within nature as well within a human designed computer. While both the 'genomic computer' and a digital computer can perform logical calculations, there are some fundamental differences. As already mentioned, cells are better in exploiting the physico-chemical properties, silicon chips rely on multiple layer of abstractness. Information passes within a digital computer in the form of flow of electricity, cells achieve this through the change in electric potential across membranes which in turn modulates the state of activation of ion-channels (Cardelli, 2005).

Another major advantage of the 'genomic computer' is the absence of hardware-software barrier. In case of traditional computers, hardwares are typical electronic units which can only function when loaded with certain softwares or logically calculating features. While in a cellular system, the DNA itself serves as information storage material, a component of expression regulatory complex and many other purposes, most important of which is providing the necessary protein coding sequence. It can be compared with an imaginary digital computer where the processor itself stores all the required information and executes certain sets of programs depending upon the requirement. Also this imaginary processor is capable of re-writing its logics based on its experience as well as capable of creating another processor to which it can pass its essential and favored characteristics and logics.

Mechanical and metabolic obligations of a cell are taken care of by the biochemical networks which are mainly the interplay of proteins. At any given time point there exist several thousand molecules of proteins and there interaction network is tremendously complex which makes it really hard to represent protein interaction network in a logical manner. Especially the context dependent modifications of proteins and altered nature in presence of other proteins or factors make it extremely hard to model protein interaction networks. Kohn maps (Kohn, 1999) and textual bio-calculus (Nagasaki, Onami, Miyano, & Kitano, 1999) have made significant contributions in this aspect; however they are far from the desired level.

As already mentioned, engineering of biological cells for the purpose of computation, is another important aspect of cellular computing. Pioneering works by L. Adleman showed that biological macromolecules like DNA can be employed to solve mathematical works by utilizing their information storage and processing capabilities (Adleman, 1994). Since then considerable progress has been made reporting the uses of DNA, enzymes and other cellular macromolecules for solving computational and mathematical challenges. However engineering of the entire biological cells for such purposes is still at its infancy. Synthetic biology deals with this problem with an aim of building an artificial cell or artificial biological system for better understanding of natural phenomena.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Synthetic biology also incorporates modular architecture like building of bio-modules like pulse generator circuits which display responses (although transient in nature) to various input stimulus (Andrianantoandro, Basu, Karig, & Weiss, 2006). Due to inherent modularity, synthetic biology can also be compared with computer engineering (Endy, 2005).

Several approaches have been taken to artificially constitute a cell that only contains the essential number of genes that can produce proteins to undertake all the necessary biochemical pathways to support life, at least at bacterial level. Szostak and group (Sazani, Larralde, & Szostak, 2004) reported the creation of an artificial nucleotide set that is self-replicating in nature (Murtas, Kuruma, Bianchini, Diaspro, & Luisi, 2007; Paul, Springsteen, & Joyce, 2006).

Beyond synthetic and systemic approach to understand biological systems, '*in vivo* computation' or 'computation within a living cell' is the most direct approach to use cells for computing purposes. Approaches have already been undertaken to develop an '*in vivo* programming' and 'autonomous finite-state automation' within the *E.coli* bacteria (Nakagawa, Sakamoto, & Sakakibara, 2006). Construction of *in-vivo* logic gates as well as genetic circuits with ability to harness the (logical) power of cellular biochemical network have been reported (Ehrenfeucht, Harju, & Petre, 2004; Landweber & Kari, 1999). With insights gained from bacterial recombination and ciliate genome transfer, scientists working in the area of '*in vivo* computing' came up with the idea of 'bioware' which is based on template guided genetic recombination (Angeleska, Jonoska, Saito, & Landweber, 2007; Nowacki, Shetty, & Landweber, 2011).

CC AND BIOINFORMATICS – IMPLEMENTATION AND OPPORTUNITES: A CASE REPORT

Bioinformatics, especially the underlying algorithms of its associated tools, has immensely benefited from CC, especially from bio-inspired algorithms (Aboul-Ella Hassanien, Mariofanna G. Milanova & Abraham, 2008; Baldi & Brunak, 2001; Hassanien, Al-Shammari, & Ghali, 2013; S. Li, Kang, & Zhao, 2014; Mulder, Akinola, Mazandu, & Rapanoel, 2014; Valentini, Tagliaferri, & Masulli). This topic has been discussed several times. In this section, the authors describe a case report on characterization of a specific protein and its interaction that have taken place with the help of bioinformatics.

Cellular actin cytoskeleton is dynamically regulated by a number of proteins including formins which are further fine-tuned by formin-binding proteins. Among several formin binding proteins, one such protein namely formin binding protein 4 (FNBP4) was discovered way back in 1996 (Bedford, Sarbassova, Xu, Leder, & Yaffe, 2000; Chan, Bedford, & Leder, 1996) as a formin poly-proline rich region (PRR) interacting protein. However in the absence of further experimental characterization, further insights in to this FNBP4 - formin PRR interaction was never reported. Even some authors later doubted the practical feasibility of this interaction (Kondo et al., 2013). In this context several bioinformatic tools like BLAST (Altschul, Gish, Miller, Myers, & Lipman, 1990), Pfam domain search (Punta et al., 2012), molecular modeling and docking techniques, many of which rely on bio-inspired algorithms for search optimization, better output generation played significant role to computationally characterize this protein and its interactions with formin (Das, Bhattacharya, Bagchi, & Dasgupta, 2013, 2014, 2015a, 2015b), not only in human but also across several other organisms across different taxonomic groups. While several characteristics of FNBP4 were discovered through bioinformatics, there are numerous other properties of this protein is yet to be discovered. This protein is mostly reported to be a part of eukaryotic multi-protein interaction network. So discovery of its interaction network will prove to be very significant and

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CC, especially the use of bio-inspired algorithms can play important role in this aspect. On the other hand FNBP4 being a large protein (average length of 800-1100 amino acids) (Das et al., 2014) would prove to be significantly challenging to determine its structure. Molecular modeling using rough sets data, especially using only amino acids, with the help of optimization algorithms also seems to be highly promising to address this problem. Thus it can be envisioned that implementation and opportunities of CC are far beyond than currently applied.

FUTURE CHALLENGES OF CELLULAR COMPUTING

Characteristics of CC lie in the interplay between the properties of the model and its implementation potentials. On should consider both of these aspects when trying to implement CC to solve some problem. Otherwise, which is often the case; the limitations of the underlying model would have serious effects on its implementations.

While simulations, mainly software based simulations have played the major role among experiments with CC, the ultimate challenge is to develop and successfully build a system which is based on several properties of CC. As highlighted by D. Hillis (Hillis, 1985), the major advantage of CC is the extreme lower level of connectivity which significantly lowers the number and cumulative bus(s) length and bus associated delay.

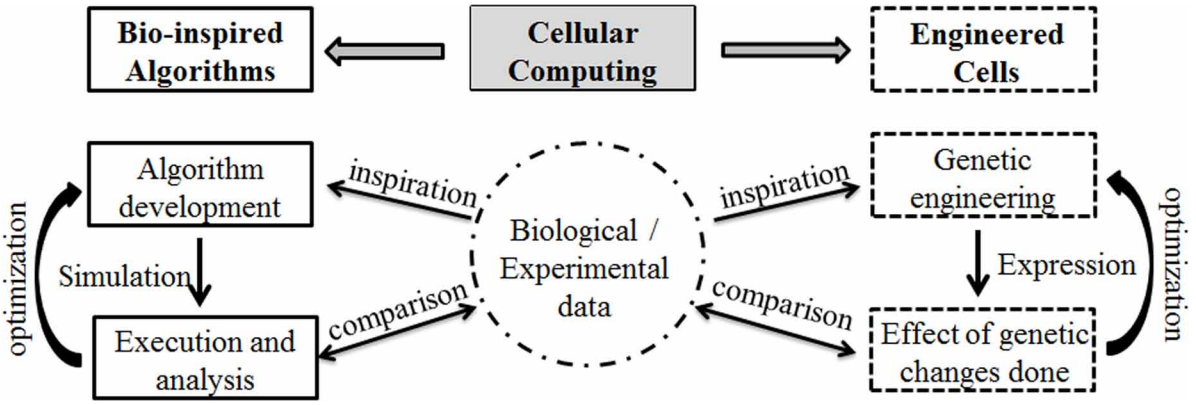
On the other hand one important fact that should always be kept in mind that biological systems or cells are a continuous system. Till date artificially inducing life in to biological components has not been achieved i.e. if all the components of a cell are assembled together, they do not necessary start to function as a living cell. A living cell can only originate from a living cell and continue to perform in similar manner till its underlying biochemical pathways continue to function in a coordinated manner. Once this coordination is lost, there is no way to get it back. Although the several advantages of a cell that we have discussing throughout this chapter, this limitation is still one of the biggest hurdle to overcome by scientists. Synthetic biology, with inputs provided by systems biology and CC models should, should apply greater focus in crossing this hurdle, at least for a unicellular or simple bacterial cells.

As highlighted in Figure 3, both arms of CC should work hand to hand and in synchronization with data provided by experimental biology as well as systems biology. While development and optimization of bio-inspired algorithms are dependent on training data set provided by classical biological observations, simulated executions can provide significant information for experimental biologists which in turn results in faster achievements of results and increased saving of time and physical resources. The field of bioinformatics is playing a greater role now a day. Several tools and techniques that are routinely employed in the field of bioinformatics rely on EA or similar algorithms for optimization purposes. Modeling of disease networks are increasingly helping in computer aided drug discovery (CADD) and over the last few decades several drugs have been designed based on inputs gained from CADD related studies. Not only drug design, molecular dynamics simulations are gaining more and more acceptance among biologists to gain detailed time to timer information about protein-protein interactions, protein-DNA and protein-small molecule interactions. Several biological interactions are explored with deeper insights (Das et al., 2013, 2015a, 2015b) with the help of bioinformatics.

It is very important to realize that CC is not an isolated approach with two aims – i) development of bio-inspired algorithms and ii) engineering of cells. Rather CC should be view as an integrated development platform which incorporates the power of computer science, experimental biology, systems biology,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Schematic diagram showing the ways biological (experimental) data is used for the purpose of cellular computing, both for algorithm development and engineered cells
First of all inspirations are drawn from biological data (both traditionally known facts and experimentally obtained) for either algorithm development (in case of bio-inspired algorithms) or genetic engineering (engineered cells). Thereafter they are either simulated or expressed, depending upon the type, and the output is compared with the already known biological data. Based on the comparison of expected output, known data and achieved output, further optimizations are undertaken.



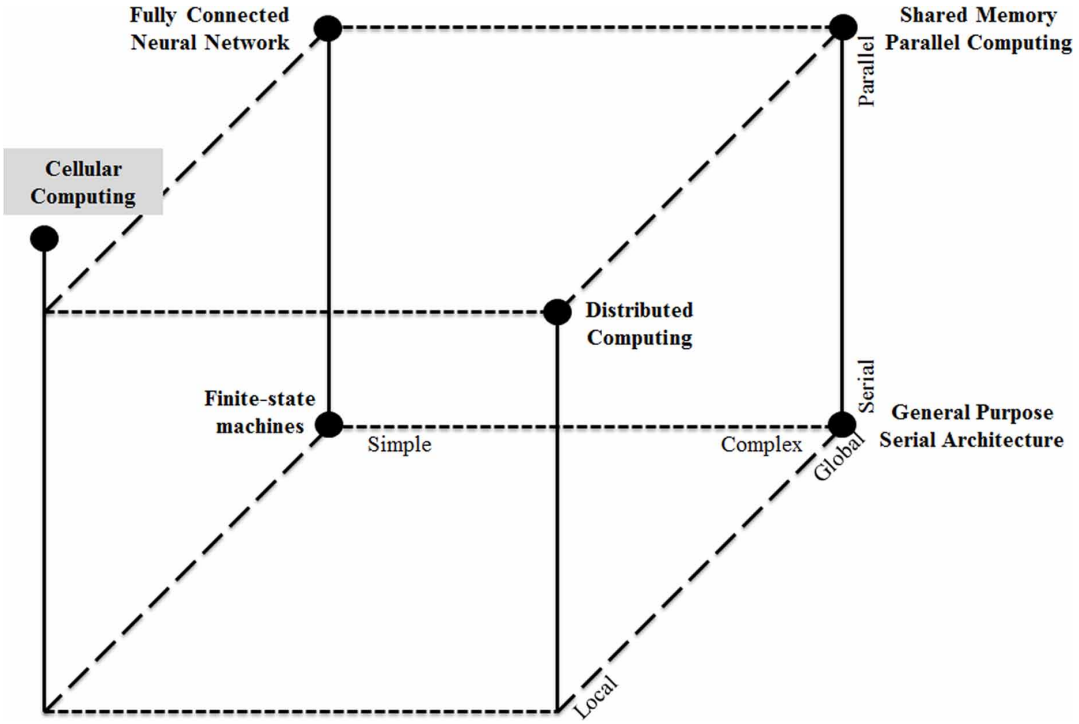
bioinformatics and computational biology in order to achieve greater benefit to the entire society. While such combined approach will provide the power the digital computer to solve biological challenges with the use of efficient algorithms, it will also help to reach computer beyond the limitations of von Neumann architecture and Moore's law (Kish, 2002; Moore, 1965; Sipper, 1999).

CONCLUSION

Simplicity, Parallelization and locality are the three main principle of CC (Figure 4). Combined with the in-built robustness (presence of only YES/NO type of outcome) and scalability (due to the absence of any central processor with which every node is required to be connected in a traditional computer system and presence of individual homogeneous nodes), implementation is fairly simple. Here comes the biggest question about areas of implementation i.e. the challenges where CC possesses inherent advantages compared to the traditional silicon chips. It can be tempting to envision biological entities performing standard mathematical or statistical calculations or bio-macromolecules storing data, but it should be noted that these are the primary targets or aims of molecular and quantum computing where an inspirational level of success have already been achieved. However specifically with CC, achievement of certain other sets of goals seems more logical and could potentially cater the urgent need of present day human civilization. Thus the aims or CC or the areas of implementation of CC can be broadly classified to three groups – A) Exploitation of inherent or embedded computing capabilities of a cell for diagnostic or therapeutic purposes, B) Using biological entities for advanced sensing purposes and thereby facilitating early or on time estimation or detection of probable or up-coming or ongoing changes, especially the ones associated with environment, C) breaking barriers of complexity and technological limits associated with silicon chips to develop advanced and efficient (energy and other costs) systems.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Properties of cellular computing
The three important aspect of cellular computing i.e. simple, parallel and local are highlighted in comparison to different properties of several other computing platforms (or architectures).



ACKNOWLEDGMENT

AD would like to thank Simanti Bhattacharya (Kalyani University, India) for critically reviewing the manuscript.

REFERENCES

Adleman, L. M. (1994). Molecular computation of solutions to combinatorial problems. *Science*, *266*(5187), 1021–1024. doi:10.1126cience.7973651 PMID:7973651

Afek, Y., Alon, N., Barad, O., Hornstein, E., Barkai, N., & Bar-Joseph, Z. (2011). A biological solution to a fundamental distributed computing problem. *Science*, *331*(6014), 183–185. doi:10.1126cience.1193210 PMID:21233379

Alon, N., Babai, L., & Itai, A. (1986). A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, *7*(4), 567–583. doi:10.1016/0196-6774(86)90019-2

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. doi:10.1016/S0022-2836(05)80360-2 PMID:2231712

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Andrianantoandro, E., Basu, S., Karig, D. K., & Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology, 2*. doi:10.1038/msb4100073

Angeleska, A., Jonoska, N., Saito, M., & Landweber, L. F. (2007). RNA-guided DNA assembly. *Journal of Theoretical Biology*, *248*(4), 706–720. doi:10.1016/j.jtbi.2007.06.007 PMID:17669433

Arbib, M. A. (1995). *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: MIT Press.

Azizi, R. (2014). Empirical Study of Artificial Fish Swarm Algorithm. *International Journal of Computing*. *Communications and Networking*, *3*(1), 1–7.

Back, T. (1996). *Evolutionary Algorithms in Theory and Practice*. Oxford University Press.

Balaji, S., Iyer, L. M., Aravind, L., & Babu, M. M. (2006). Uncovering a hidden distributed architecture behind scale-free transcriptional regulatory networks. *Journal of Molecular Biology*, *360*(1), 204–212. doi:10.1016/j.jmb.2006.04.026 PMID:16730024

Baldi, P., & Brunak, S. (2001). *Bioinformatics: the machine learning approach*. Cambridge, MA: MIT Press.

Bedford, M. T., Sarbassova, D., Xu, J., Leder, P., & Yaffe, M. B. (2000). A novel pro-Arg motif recognized by WW domains. *The Journal of Biological Chemistry*, *275*(14), 10359–10369. doi:10.1074/jbc.275.14.10359 PMID:10744724

Benenson, Y. (2012). Biomolecular computing systems: Principles, progress and potential. *Nature Reviews. Genetics*, *13*(7), 455–468. doi:10.1038/nrg3197 PMID:22688678

Binitha, S., & Sathya, S. S. (2012). A Survey of Bio inspired Optimization Algorithms. *International Journal of Soft Computing and Engineering*, *2*(2), 137–151.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press.

Blum, C. (2005). Ant colony optimization: Introduction and recent trends. *Physics of Life Reviews*, *2*(4), 353–373. doi:10.1016/j.plrev.2005.10.001

Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm Intelligence*. Oxford University Press.

Cardelli, L. (2005). Brane calculi: Interactions of biological membranes. In V. Danos & V. Schachter (Eds.), Computational Methods in Systems Biology (LNCS), (vol. 3082, pp. 257–278). Berlin: Springer Berlin Heidelberg; doi:10.1007/978-3-540-25974-9_24

Cardelli, L. (2007). Machines of systems biology. *Bulletin of the EATCS*, *93*, 176–204.

Chan, D. C., Bedford, M. T., & Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *The EMBO Journal*, *15*(5), 1045–1054. PMID:8605874

Chen, H., & Zhu, Y. (2008). Optimization based on symbiotic multi-species coevolution. *Applied Mathematics and Computation*, *205*(1), 47–60. doi:10.1016/j.amc.2008.05.148

Conrad, M. (1989). The brain-machine disanalogy. *Bio Systems*, *22*(3), 197–213. doi:10.1016/0303-2647(89)90061-0 PMID:2650754

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Das, A., Bhattacharya, S., Bagchi, A., & Dasgupta, R. (2013). Structural insight into the formin binding property of human formin binding protein 4. *Journal of Proeins and Proteomics*, *4*(December), 93–100.

Das, A., Bhattacharya, S., Bagchi, A., & Dasgupta, R. (2014). In-silico characterization of Formin Binding Protein 4 Family of proteins. *Interdisciplinary Sciences, Computational. Life Sciences*. doi:10.100712539-013-0040-9

Das, A., Bhattacharya, S., Bagchi, A., & Dasgupta, R. (2015a). Structural and functional characterization of Arabidopsis thaliana WW domain containing protein F4JC80. In Computational Intelligence in Medical Informatics (pp. 25–39). doi:10.1007/978-981-287-260-9_3

Das, A., Bhattacharya, S., Bagchi, A., & Dasgupta, R. (2015b). Understanding the Interaction of Human Formin Binding Protein 4 with Formin FMN1. In Information Systems Design and Intelligent Applications (pp. 105–113). doi:10.1007/978-81-322-2247-7_12

Dasgupta, D. (1999). *Artificial Immune Systems and Their Applications*. Springer Berlin. doi:10.1007/978-3-642-59901-9

Davidson, E., & Levin, M. (2005). Gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(14), 4935. doi:10.1073/pnas.0502024102 PMID:15809445

Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics : A Publication of the IEEE Systems. Man, and Cybernetics Society*, *26*(1), 29–41. doi:10.1109/3477.484436

Dorigo, M., & Stützle, T. (2004). *Ant Colony Optimization*. Bradford Company. doi:10.1007/b99492

Dorigo, M., & Stützle, T. (2010). Ant Colony Optimization: Overview and Recent Advances. In Handbook of Metaheuristics (pp. 227–263). Academic Press.

Ehrenfeucht, A., Harju, T., & Petre, I. (2004). *Computation in Living Cells: Gene Assembly in Ciliates*. Academic Press.

Endy, D. (2005). Foundations for engineering biology. *Nature*, *438*(7067), 449–453. doi:10.1038/nature04342 PMID:16306983

Fe-Fei, L., Fergus, R., & Perona, P. (2003). A Bayesian approach to unsupervised one-shot learning of object categories. In *Proceedings Ninth IEEE International Conference on Computer Vision* (p. 1134). IEEE Computer Society. 10.1109/ICCV.2003.1238476

Fei-Fei, L., Fergus, R., & Perona, P. (2006). One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(4), 594–611. doi:10.1109/TPAMI.2006.79 PMID:16566508

Feynman, R. P. (1960). There's plenty of room at the bottom: An invitation to enter a new field of physics. In Caltech's Eng Sci (pp. 22–36). doi:10.1109/84.128057

Fisher, J., & Henzinger, T. (2007). Executable cell biology. *Nature Biotechnology*, *25*(11), 1239–1249. doi:10.1038/nbt1356 PMID:17989686

Fu, P. (2007). Biomolecular computing: Is it ready to take off? *Biotechnology Journal*, *2*(1), 91–101. doi:10.1002/biot.200600134 PMID:17183505

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Girardini, J. E., Walerych, D., & Del Sal, G. (2014). Cooperation of p53 mutations with other oncogenic alterations in cancer. *Sub-Cellular Biochemistry*, *85*, 41–70. doi:10.1007/978-94-017-9211-0_3 PMID:25201188

Gitter, A., Siegfried, Z., Klutstein, M., Fornes, O., Oliva, B., Simon, I., & Bar-Joseph, Z. (2009). Backup in gene regulatory networks explains differences between binding and knockout results. *Molecular Systems Biology*, *5*, 276. doi:10.1038/msb.2009.33 PMID:19536199

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc.

Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., & Li, W.-H. (2003). Role of duplicate genes in genetic robustness against null mutations. *Nature*, *421*(6918), 63–66. doi:10.1038/nature01198 PMID:12511954

Guerraoui, R., & Schiper, A. (1996). Article. In A. Strohmeier (Ed.), Fault-tolerance by replication in distributed systems (LNCS), (Vol. 1088, pp. 38–57). Berlin: Springer Berlin Heidelberg; doi:10.1007/BFb0013474

Hassanien, Milanova, & Abraham, A. (2008). Computational Intelligence in Solving Bioinformatics Problems: Reviews, Perspectives, and Challenges. In Computational Intelligence in Biomedicine and Bioinformatics: Studies in Computational Intelligence (pp. 3–47). Springer Berlin Heidelberg. doi:10.1007/978-3-540-70778-3_1

Hassanien, A. E., Al-Shammari, E. T., & Ghali, N. I. (2013). Computational intelligence techniques in bioinformatics. *Computational Biology and Chemistry*, *47*, 37–47. doi:10.1016/j.compbiolchem.2013.04.007 PMID:23891719

Hecker, M., Lambeck, S., Toepfer, S., van Someren, E., & Guthke, R. (2009). Gene regulatory network inference: Data integration in dynamic models-a review. *Bio Systems*, *96*(1), 86–103. doi:10.1016/j.biosystems.2008.12.004 PMID:19150482

Hillis, W. D. (1985). *The Connection Machine*. Cambridge, MA: MIT Press.

Hooke, R. (1665). *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses: With observations and inquiries thereupon /by R. Hooke ...*. London: Royal Society. doi:10.5962/bhl.title.904

Istrail, S., De-Leon, S. B.-T., & Davidson, E. H. (2007). The regulatory genome and the computer. *Developmental Biology*, *310*(2), 187–195. doi:10.1016/j.ydbio.2007.08.009 PMID:17822690

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, *29*(3), 31–44. doi:10.1109/2.485891

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks* (Vol. 4, pp. 1942–1948). IEEE. 10.1109/ICNN.1995.488968

Kish, L. B. (2002). End of Moore's law: Thermal (noise) death of integration in micro and nano electronics. *Physics Letters. [Part A]*, *305*(3-4), 144–149. doi:10.1016/S0375-9601(02)01365-8

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kitano, H. (2001). *Fondations of Systems Biology*. Cambridge, MA: MIT Press.

Kohn, K. W. (1999). Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Molecular Biology of the Cell*, *10*(8), 2703–2734. doi:10.1091/mbc.10.8.2703 PMID:10436023

Kondo, Y., Koshimizu, E., Megarbane, A., Hamanoue, H., Okada, I., Nishiyama, K., ... Matsumoto, N. (2013). Whole-exome sequencing identified a homozygous FNBP4 mutation in a family with a condition similar to microphthalmia with limb anomalies. *American Journal of Medical Genetics. Part A*, *161A*(7), 1543–1546. doi:10.1002/ajmg.a.35983 PMID:23703728

Kuznetsov, V., Lee, H. K., Maurer-Stroh, S., Molnár, M. J., Pongor, S., Eisenhaber, B., & Eisenhaber, F. (2013). How bioinformatics influences health informatics: Usage of biomolecular sequences, expression profiles and automated microscopic image analyses for clinical needs and public health. *Health Information Science and Systems*, *1*(1), 2. doi:10.1186/2047-2501-1-2 PMID:25825654

Lai, Y.-H., Sun, S.-C., & Chuang, M.-C. (2014). Biosensors with built-in biomolecular logic gates for practical applications. *Biosensors*, *4*(3), 273–300. doi:10.3390/bios4030273 PMID:25587423

Landweber, L. F., & Kari, L. (1999). The evolution of cellular computing: Nature's solution to a computational problem. *Bio Systems*, *52*(1-3), 3–13. doi:10.1016/S0303-2647(99)00027-1 PMID:10636025

Latty, T., & Beekman, M. (2011). Speed-accuracy trade-offs during foraging decisions in the acellular slime mould Physarum polycephalum. *Proceedings. Biological Sciences / The Royal Society, 278*(1705), 539–45. doi:10.1098/rspb.2010.1624

Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., ... Young, R. A. (2002). Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science*, *298*(5594), 799–804. doi:10.1126cience.1075090 PMID:12399584

Li, S., Kang, L., & Zhao, X.-M. (2014). A survey on evolutionary algorithm based hybrid intelligence in bioinformatics. *BioMed Research International*, *362738*. doi:10.1155/2014/362738 PMID:24729969

Li, X., & Shao, Z. (2002). An optimizing method based on autonomous animales: Fish swarm algorithm. *Systems Engineering Theory and Practice*, *22*, 32–38.

Luby, M. (1985). A simple parallel algorithm for the maximal independent set problem. In *Proceedings of the seventeenth annual ACM symposium on Theory of computing* (pp. 1–10). New York: ACM Press. 10.1145/22145.22146

Marwan, W. (2010). Systems biology. Amoeba-inspired network design. *Science*, *327*(5964), 419–420. doi:10.1126cience.1185570 PMID:20093462

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133. doi:10.1007/BF02478259

Mehrabian, A. R., & Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization. *Ecological Informatics*, *1*(4), 355–366. doi:10.1016/j.ecoinf.2006.07.003

Moore, G. E. (1965). Cramming more components onto integrated circuits (Reprinted from Electronics, pg 114-117, April 19, 1965). *Electronics*, *38*(8), 114–117. doi:10.1109/N-SSC.2006.4785860

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Mulder, N. J., Akinola, R. O., Mazandu, G. K., & Rapanoel, H. (2014). Using biological networks to improve our understanding of infectious diseases. *Computational and Structural Biotechnology Journal*, *11*(18), 1–10. doi:10.1016/j.csbj.2014.08.006 PMID:25379138

Muller, P. A. J., & Vousden, K. H. (2014). Mutant p53 in cancer: New functions and therapeutic opportunities. *Cancer Cell*, *25*(3), 304–317. doi:10.1016/j.ccr.2014.01.021 PMID:24651012

Murtas, G., Kuruma, Y., Bianchini, P., Diaspro, A., & Luisi, P. L. (2007). Protein synthesis in liposomes with a minimal set of enzymes. *Biochemical and Biophysical Research Communications*, *363*(1), 12–17. doi:10.1016/j.bbrc.2007.07.201 PMID:17850764

Nagasaki, M., Onami, S., Miyano, S., & Kitano, H. (1999). Bio-calculus: Its Concept and Molecular Interaction. *Genome Informatics. Workshop on Genome Informatics*, *10*, 133–143.

Nakagawa, H., Sakamoto, K., & Sakakibara, Y. (2006). Development of an in vivo computer based on escherichia coli. In A. Carbone & N. A. Pierce (Eds.), *DNA Computing* (Vol. 3892, pp. 203–212). Berlin: Springer Berlin Heidelberg; doi:10.1007/11753681

Navlakha, S., & Bar-Joseph, Z. (2011). Algorithms in nature: The convergence of systems biology and computational thinking. *Molecular Systems Biology*, *7*(546), 1–11. doi:10.1038/msb.2011.78 PMID:22068329

Nowacki, M., Shetty, K., & Landweber, L. F. (2011). RNA-Mediated Epigenetic Programming of Genome Rearrangements. *Annual Review of Genomics and Human Genetics*, *12*(1), 367–389. doi:10.1146/annurev-genom-082410-101420 PMID:21801022

Pace, B. S., Ofori-Acquah, S. F., & Peterson, K. R. (2012). Sickle cell disease: Genetics, cellular and molecular mechanisms, and therapies. *Anemia*, *2012–2014*. doi:10.1155/2012/143594 PMID:22957221

Papamichos-Chronakis, M., & Peterson, C. L. (2013). Chromatin and the genome integrity network. *Nature Reviews. Genetics*, *14*(1), 62–75. doi:10.1038/nrg3345 PMID:23247436

Pattee, H. H. (1961). On the origin of macromolecular sequences. *Biophysical Journal*, *1*(8), 683–710. doi:10.1016/S0006-3495(61)86917-8 PMID:14484391

Paul, N., Springsteen, G., & Joyce, G. F. (2006). Conversion of a ribozyme to a deoxyribozyme through in vitro evolution. *Chemistry & Biology*, *13*(3), 329–338. doi:10.1016/j.chembiol.2006.01.007 PMID:16638538

Powell, E., Piwnica-Worms, D., & Piwnica-Worms, H. (2014). Contribution of p53 to metastasis. *Cancer Discovery*, *4*(4), 405–414. doi:10.1158/2159-8290.CD-13-0136 PMID:24658082

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., ... Finn, R. D. (2012). The Pfam protein families database. *Nucleic Acids Research*, *40*(Database issue), D290–D301. doi:10.1093/nar/gkr1065 PMID:22127870

Rendel, J. M., Sheldon, B. L., & Finlay, D. E. (1965). Canalisation of development of scutellar bristles in Drosophila by control of the scute locus. *Genetics*, *52*(6), 1137–1151. PMID:5882192

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sazani, P. L., Larralde, R., & Szostak, J. W. (2004). A small aptamer with strong and specific recognition of the triphosphate of ATP. *Journal of the American Chemical Society*, *126*(27), 8370–8371. doi:10.1021/ja049171k PMID:15237981

Schmitt, F. O. (1962). *Macromolecular Specificity and Biological Memory*. Cambridge, MA: MIT Press.

Shah-Hosseini, H. (2007). Problem solving by intelligent water drops. In *2007 IEEE Congress on Evolutionary Computation* (pp. 3226–3231). IEEE. 10.1109/CEC.2007.4424885

Shah-Hosseini, H. (2008). Intelligent water drops algorithm: A new optimization method for solving the multiple knapsack problem. *International Journal of Intelligent Computing and Cybernetics*, *1*(2), 193–212. doi:10.1108/17563780810874717

Simon, D. (2008). Biogeography-Based Optimization. *IEEE Transactions on Evolutionary Computation*, *12*(6), 702–713. doi:10.1109/TEVC.2008.919004

Sipper, M. (1999). The emergence of cellular computing. *Computer*, *32*(7), 18–26. doi:10.1109/2.774914

Sirbu, B. M., & Cortez, D. (2013). DNA damage response: Three levels of DNA repair regulation. *Cold Spring Harbor Perspectives in Biology*, *5*(8), a012724. doi:10.1101/cshperspect.a012724 PMID:23813586

Stahl, W. R., & Goheen, H. E. (1963). Molecular algorithms. *Journal of Theoretical Biology*, *5*(2), 266–287. doi:10.1016/0022-5193(63)90064-X PMID:5875393

Sugita, M. (1961). Functional analysis of chemical systems in vivo using a logical circuit equivalent. *Journal of Theoretical Biology*, *1*, 415–430. PMID:13918223

Sugita, M. (1963). Functional analysis of chemical systems in vivo using a logical circuit equivalent. II. The idea of a molecular automation. *Journal of Theoretical Biology*, *4*(2), 179–192. doi:10.1016/0022-5193(63)90027-4 PMID:5875160

Sugita, M., & Fukuda, N. (1963). Functional analysis of chemical systems in vivo using a logical circuit equivalent. 3. Analysis using a digital circuit combined with an analogue computer. *Journal of Theoretical Biology*, *5*(3), 412–425. doi:10.1016/0022-5193(63)90087-0 PMID:5336375

Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D. P., Fricker, M. D., ... Nakagaki, T. (2010). Rules for biologically inspired adaptive network design. *Science*, *327*(5964), 439–442. doi:10.1126cience.1177894 PMID:20093467

Teuscher, C. (2002). *Turing's Connectionism - An Investigation of Neural Network Architectures*. Springer London.

Teuscher, C. (2012). Cellular computing. In Encyclopedia of Complexity and Systems Science, (pp. 465–478). Academic Press.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. doi:10.1038/381520a0 PMID:8632824

Turing, A. M. (1969). Intelligent machinery. In B. Meltzer & D. Michie (Eds.), *Machine Intelligence* (pp. 3–23). Edinburgh, UK: Edinburgh University Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Valentini, G., Tagliaferri, R., & Masulli, F. (n.d.). Computational intelligence and machine learning in bioinformatics. *Artificial Intelligence in Medicine, 45*(2-3), 91–6. doi:10.1016/j.artmed.2008.08.014

Watanabe, S., Tero, A., Takamatsu, A., & Nakagaki, T. (2011). Traffic optimization in railroad networks using an algorithm mimicking an amoeba-like organism, Physarum plasmodium. *Bio Systems*, *105*(3), 225–232. doi:10.1016/j.biosystems.2011.05.001 PMID:21620930

Yeudall, W. A. (2014). p53 mutation in the genesis of metastasis. *Sub-Cellular Biochemistry*, *85*, 105–117. doi:10.1007/978-94-017-9211-0_6 PMID:25201191

Yip, V., Beekman, M., & Latty, T. (2014). Foraging strategies of the acellular slime moulds Didymium iridis and Didymium bahiense. *Fungal Ecology*, *11*, 29–36. doi:10.1016/j.funeco.2014.02.001

Yovel, Y., Falk, B., Moss, C. F., & Ulanovsky, N. (2010). Optimal localization by pointing off axis. *Science*, *327*(5966), 701–704. doi:10.1126cience.1183310 PMID:20133574

Zhu, X., Rogers, T., Qian, R., & Kalish, C. (2007). Humans Perform Semi-supervised Classification Too. In *Proceedings of the 22Nd National Conference on Artificial Intelligence (AAAI'07)* (pp. 864–869). AAAI Press.

KEY TERMS AND DEFINITIONS

Bio-Inspired Algorithms: Computational algorithms which has been designed and optimized with inputs from the behavior of biological systems like an ant or bee colony or biological entities like a cell or neuron. Example: artificial neural network, genetic algorithm, membrane computing etc.

Cell: A cell is an independently sustainable and self-replicating unit of any organism. For multi-cellular organisms, a cell also serves as the structural and functional unit. Cells bear the capability of responding to surrounding environment through logical outputs of the received external or internal signals.

Central Dogma: The flow of genetic information, coded in DNA, to proteins via RNA is termed as the central dogma of molecular biology. It is mainly composed of two processes i.e. transcription which synthesizes RNA from DNA and translation which converts the RNA encoded message into a string of amino acids that subsequently folds in to a protein.

Genetic Engineering: A field of molecular biology that deals with purposeful manipulation of an organisms (or cells) native genetic content for the gain or deletion of certain traits through addition of foreign DNA material or deletion of certain portion of native DNA.

Modeling of Biological Systems: A part of systems biology that deals with the computational modeling of biological systems or pathways like signaling cascade, metabolic network, disease network etc. The main purpose of such modeling is simulated replications of biological phenomena for ease of understanding in the event of deviations of certain parameters from natural values and prediction of their possible impact.

Molecular Computation: The use of biological macro-molecules like DNA or proteins (like enzymes) for the purpose of computation like calculations, storing, analyses or processing of data is known as molecular computation or bio-molecular computation.

Natural Computing: Natural computing is an interdisciplinary arena that primarily encompasses computer science and biology along with indirect association with all other possible scientific fields.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Natural computing primarily encompasses three major areas – A) Development and optimizations of computational algorithms with inspirations from nature, mainly biological systems, B) Computation modeling of biological systems or pathways and C) Use of biological entities like cells for the purpose of computation.

Parallel Computing: A type of computation which facilitates simultaneous execution of different tasks which are primarily sub-tasks of a larger task. The main aim of parallel computing is improvement of total execution time by distributing parts of a task to separate computing nodes or cores from a single such computing unit.

Systems Biology: A biology centered interdisciplinary field of study that focuses on systemic study (holistic approach) of complex interactions of biological systems. With insights gained from experimental biology, systems biology incorporates computational and mathematical (with some statistical inputs) modeling and simulation techniques and tools for a comprehensive and quantitative analysis of biological interactions.

This research was previously published in the Handbook of Research on Natural Computing for Optimization Problems edited by Jyotsna Kumar Mandal, Somnath Mukhopadhyay, and Tandra Pal; pages 637-662; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 96
From Telecytology to Mobile Cytopathology: Past, Present, and Future

Abraham Pouliakis
National and Kapodistrian University of Athens, Greece

Stavros Archondakis
401 Military Hospital, Greece

Niki Margari
National and Kapodistrian University of Athens, Greece

Petros Karakitsos
National and Kapodistrian University of Athens, Greece

ABSTRACT

Cytopathology is a popular discipline since George Papanicolaou proposed the famous test pap about 60 years ago. Today modern cytopathology laboratories still use the microscope as the primary diagnostic device and additionally they host modalities performing medical tests and exchange data via networks and have imaging systems producing pictures and virtual slides; the volume of produced data nowadays is enormous. Simultaneously mobile phones and tablets have evolved; their capabilities compete desktop computers and have the advantage of being always connected and at the side of users. Despite there are rather limited applications relevant to cytopathology for the mobile device, there is potential for uses in numerous activities of the cytopathology laboratory, including and not limited to: training, reporting, diagnosis and consultation, laboratory management, whole slide imaging, interactions between patient-doctor, doctor-doctor and within the laboratory personnel, quality control and assurance. Mobile devices can offer important benefits to the modern cytopathology laboratory.

DOI: 10.4018/978-1-7998-1204-3.ch096

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Mobile Health (mHealth) is the practice of medicine and public health supported by mobile devices. The use of mobile computing and communication devices, such as mobile phones, tablet computers and personal digital assistants (PDAs), by health professionals, is nowadays rapidly growing. The mHealth applications are mainly used for collecting community and clinical health data, delivery of healthcare information, real-time monitoring of patient vital signs, and direct healthcare provisioning. In more detail, handheld computing has applications such as ambulatory medicine (K. Banitsas et al., 2005; K. A. Banitsas et al., 2006; Kiselev et al., 2012; Pavlopoulos et al., 1998; Rosales Saurer et al., 2009; Zerth et al., 2012), diabetes management (Quinn et al., 2011; Ribu et al., 2013; Skrovseth et al., 2012; Spat et al., 2013), asthma management (Finkelstein et al., 1998; Gupta et al., 2011), control of obesity (Patrick et al., 2009), smoking control (Ghorai et al., 2014; Ybarra et al., 2014), seizure management (Pandher et al., 2014), stress management (Clarke et al., 2014) and treatment of depression (Burns et al., 2011) among others. In the field of mHealth the majority of applications are for fitness (43%) followed by health resource (15.0%) and diet/caloric intake (14.3%,), while the user angagement has the form of self-monitoring and training (74.8%) (Sama et al., 2014); in contrast, despite there are applications targeting patinets, currently, there are rather limited applications targeting physicians and doctor-patient interactions (T. Martin, 2012). Pioneering field seems to be radiology consultation for X-rays and mostly Computer Tomography (Choudhri et al., 2012; Choudhri et al., 2013; Johnson et al., 2012; Toomey et al., 2010) and ECG transmission (Vaisanen et al., 2003). The applications are very limited in the fields of pathology and even less in cytopathology despite both specialties deal with images. In relation to pathology, the most reported uses of handhelds, are limited to experimental endeavors in education and telemedicine (Park et al., 2012). For cytopathology, even after a thorough search, there were not found published articles or reports. However, pathology and cytopathology share many common characteristics. Actually, in most countries, cytopathology is considered as a subspecialty of pathology. Thus concepts and ideas can be useful to both specialties; therefore applications can be transferred from one domain to the other.

Cytopathology is a medical sector/discipline that gained popularity when George Papanicolaou proposed the famous test Papanicolaou (known as Pap Test), about 60 years ago. Nowadays Pap Test is the most valuable tool for cervical cancer screening and prevention. Therefore, it is the reason that cytopathology is so popular. Additionally, cytopathology has the advantage of obtaining biological material using minimally invasive or not invasive at all methods. Cytopathology is a discipline that the diagnosis is mainly based on the examination of cells via the microscope, histopathology (or pathology) does the same; however using complete tissues (obtained via biopsy). The routine cytological examinations are performed, since the invention of the microscope, via the utilization of a glass slide and subsequent visual analysis. Today, the modern cytopathology laboratory is continuously changing. Nowadays, cytopathology laboratories perform additional examinations based on molecular techniques and immunocytochemistry methods. The modern cytopathology laboratory is equipped with a lot of modalities, these are capable of performing medical tests, as well as to exchange data via networks, there are available as well imaging systems, capable to create digital pictures of the slides or even virtual slides, which are complete slides in electronic format. The volume of data, in a cytopathology laboratory, nowadays is enormous; there are many applications that are available, and others that can be envisioned for the benefit of cytopathologists and the patients. The recent advances in handheld hardware and software; in parallel, with concurrent advances in whole slide imaging (WSI) and cloud computing, offer new opportunities

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and challenges for cytopathology. Cloud computing (one of the foundations required for mHealth) has already been proposed as a useful platform for modern cytopathology (Pouliakis et al., 2014a) as well as in other medical laboratories (Pouliakis et al., 2014d).

In this chapter we address the current state of the art in handheld hardware and software, provide the history of handheld devices, with special focus on pathology and cytopathology, and present use cases that have the potential to be future applications. Various aspects of the modern cytopathology laboratory are analyzed as well, we present a thorough research of mobile applications, related to cytopathology and try to foresee applications, that, if enhanced with mobility and are available on mobile devices, they may provide an important benefit to the modern cytopathology laboratory, as well as to the patients. Wherever possible, we propose new mobile applications, having the potential to enhance the routine of the laboratory. Finally we highlight issues, controversies and problems and try to propose solutions and recommendations.

BACKGROUND

In a two year old research (Park et al., 2012), an exhaustive search in PubMed revealed more than 6,500 articles on the topic of handheld computing, publication dates were from 1983. Of those, 55% were related to the usage of handhelds in medicine, rather than health risks caused by handheld devices (for example risk of cancer from radiation, likelihood of automobile accidents or electromagnetic interference in hospital devices). The vast majority of articles were published in the last decade; for instance, two articles were published in the year of 1983, compared to 61 articles in October 2011 alone (Park et al., 2012). A PubMed search for the terms cytology and handheld revealed 31 articles of which most are on miniaturized devices (for example mobile Flow Cytometer). However, only nine articles in the entire medical literature that deal directly with handheld computing in pathology there were found, ranging in publication date from 2004 to 2015. Based on the available literature, some conclusions may be extracted. Almost half of the scientific articles deal with static telepathology using mobile phone cameras, but there is only one paper that discusses whole slide imaging and about 25% of the articles deal with medical education. Despite not reported in the scientific medical literature, there are other innovative uses of handhelds in pathology. Nowadays, several reference laboratories, offer handheld applications that their clients can use to order tests and receive the results. There are available applications, that exploit the built-in phone camera, such as taking photographs through the eyepiece or interface a phone with a microscope directly (Breslauer et al., 2009). Concluding, the scientific literature, related to mobile pathology and cytopathology can be considered rather poor.

Perhaps there are multiple possible reasons for this. Pathology, despite not being the sole medical specialty in need to manipulate data and/or images, it is unique in the scale of data that must handle. It is considered that over 70% of the data in a typical electronic medical record are generated within the clinical pathology laboratory (Pantanowitz et al., 2007). In anatomic pathology, a Whole Slide Image (WSI) can easily be in the range of gigabytes in size, even when heavy compression with losses is applied (Pouliakis et al., 2014a); comparing this to radiology, being a medical specialty with heavy use of images as well, in which it is rather rare for a digital image to be more than a few hundred megabytes in size. In addition, radiologists work primarily on grayscale images (8-bit and now 12 or 16 bits when measurements are required) with resolutions in the thousands of pixels, pathologists work with color (24-32 bit, which, simply triples the amount of data). In pathology the resolution of images in the range

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of tens to hundreds of thousands of pixels. This dictates the need for a large amount of computational power, high-quality and high-resolution screens with large bit depth (aka color range), and fast network access (sometimes with metropolitan or global aspects and not just for within building communications). All these characteristics, which were not available in handheld (and sometimes not even on desktop) computers until very recently, were a barrier to the adoption of electronic data interchange in pathology and even worst in mobile pathology and cytopathology. Nowadays as the characteristics of mobile devices have changed, it seems that it is the time for mobile pathology (and cytopathology).

Mobile telephony is now more than 20 years in use; the recent advances, in the last decade, have contributed to the conversion of the mobile telephones into smart telephones; due to the sophisticated operating systems. Especially, during the last five years, there became available hardware components enriching mobile phones: more processing power (4 or 8 processors), high definition cameras (more than 10Mpixel nowadays) and large displays resolutions (nowadays more than 1500x2500). Moreover, the 4G and WiFi connectivity have created an always open, rather inexpensive, communication channel with the Internet. Today, there is no more the trend to shrink the telephone, in contrast, mobile phones are becoming larger and compete tablets in terms of size and weight, they are easy to use and have capabilities similar to computers available five years ago. It seems that the mobile devices have now the maturity of computers being used in health sector five years ago and are ready to be adopted. Some of the characteristics of mobile devices (most of them not relevant to the desktop computers) are summarized below:

- Battery with capacity lasting for days
- CPU with processing power competing desktop and laptops if we consider that users rarely use the full processing power
- Excellent graphics capabilities
- Processor memory in the range of gigabytes
- Flash memory for storage in the range of hundreds of Gigabytes
- Embedded GPS chips to allow global positioning
- 3G and 4G connectivity
- WiFi connectivity (150Mbit) competing standard (1,000 Mbit) LAN connectivity if we consider that LAN connectivity is shared as well
- Bluetooth connectivity for the creation of personal area networks
- USB connectivity for coupling the device to other devices
- Accelerometer
- Gyroscope
- Compass
- Thermometer
- One or more high-resolution cameras
- Large displays of high resolution
- Capacitive touch screen allowing a rich user interaction
- Connectors to interface with displays on high-end models

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

MOBILE CYTOPATHOLOGY

Over the last years, new types of cameras and microscopes, connected to computers made possible cytological image capture and transmission, named as telecytology (Archondakis et al., 2009; Pantanowitz et al., 2009; Pinco et al., 2009). The wide implementation of telemedical systems in this field became a necessity dictated by the need of real-time results for therapeutic decisions (Briscoe et al., 2000; Jialdasani et al., 2006; Mun et al., 1995; Raab et al., 1996; Weinstein et al., 1997; Yamashiro et al., 2004) as well as physician training (Stergiou et al., 2009). However, telediagnosis and training are not the sole applications relevant to cytopathology. other applications, include but are not limited to the everyday routine tasks of the modern laboratory, specifically there are interesting applications for results reporting, day to day laboratory management, Whole Slide Imaging (WSI), various interactions between patient and doctor, doctor to doctor and between members of the same laboratory, finally there are applications relevant to quality control and assurance. In the sequel we analyse the available applications and try wherever possible pinpoint potential applications.

Mobile Applications Related to Teaching

In 2006, a group of reasearchers, reported on the use of mobile phone cameras as a method of remote teaching in undergraduate pathology education (Sharma et al., 2006). Specifically, they noticed that students possessed photographs of nearly all the typical specimens they may encounter during examinations; these photos were considered by the departmental authorities harmless and similar to note-taking. Thus, they triggered a decission of the instutute to allow the mobile device during examinations, to be used as an aid to memory.

Five years later, another researcher (Collins, 2011), reported on the usage of iPads for the online distribution of digital textbooks for cytopathology. The application was based on the epub 3.0 standard, which facilitates the creation of electronic books (e-books). The author of the application used the Adobe InDesign CS5 software, designed and created an electronic book with text and images; these were possible to be accessed from a variety of different mobile platforms. The major benefits of this cytopathology e-book was the searchable content, the interactive text and references with high resolution cytopathology photomicrographs, the capability to bookmark a page for future reference after exiting, video embedding to have a multimedia rich content, instant access to medical references with a single touch of a finger. Additionally, images as well as the relevant text; could be viewed on stand-alone in full screen because the iPad could be connected to a larger, external HDTV screen. In this setting the zoom function of the iPad permitted detailed examination of image details. Moreover, content could be copied, pasted, printed, and emailed. The 64 GB capacity up the mobile device was considered as ample storage space for such applications. Nowadays, the cloud based storage and the always connected capability provides virtually unlimited storage.

Modern browsers, installed in the mobile phones today, are not proprietary anymore, in contrast they are identical to the browser installed on desktop PCs. This means that there is no need anymore, to develop web applications, specifically for the mobile device. Moreover, any advances in the usage of Web 2.0 technologies, in cytopathology related material, will be directly available to the mobile phones as well. Furthermore, handhelds today integrate numerous connectivity subsystems and surprisinglly relational databases, thus, there is the embedded capability to the handheld allowing its use both as server and client. Especially, in developing countries, and during disasters, the mobile devices have proved their

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

usefulness, mainlly because the tablets have a form factor and size similar to printed books. In 2004, another group of researchers (Ng et al., 2004), published a paper on using Internet search engines to find high-quality reference material for oral and maxillofacial pathology, and then caching this material on a PDA, nowadays caching is not a necessity as the mobile device is always connected.

Teaching applications on the mobile device, are not only for reasearch purposes. Since 2011, there are several commercial applications being placed online, the including handbooks (Elsevier, 2015), atlases, see for example the Johns Hopkins Atlas of Pancreatic Cytopathology (Meszaros, 2014) as well as the majority of cytopathology related scientific journals. Despite all these applications were designed taking to account the special characteristics of the mobile device, the modern mobile phone browsers, have turned all those pioneering efforts almost obsolete, this is because any web page can today be viewed by the mobile phone browser and the high resolution of the screen facilitates this to some degree.

Mobile Applications for Reporting the Results of the Cytopathology Laboratory

Cytopathologists and pathologists have to learn not only to formulate their diagnosis based on cellular/tissue features, but also how to communicate the results in a report being clinically useful. Indeed, even the choice of the appropriate reporting format may have consequences on patient treatment (Crowe et al., 2011; Hirsch et al., 2015). It is frequent, that medical reports produced in laboratories, do not contain all information required for clinicians, thus they are not guided towards a good choice. Moreover, these reports may be formulated inappropriately. A group of researchers (Skeate et al., 2007) in 2007, created a software environment, and subsequently reported the usage of knowledge bases on PDAs. Their main outcome was that PDAs can enhance resident learning and pathology report completion as well. Despite their application falls within the pathology discipline, similar outcomes are expected for cytopathology, because reporting for both disciplines has similar complexity. Their experimental setting included ReportSupport, a software system operational on PDAs. ReportSupport was a navigable document, developed in the Hyper Text Mark-Up Language (HTML) using Macromedia's Dreamweaver 4, and viewed using the device browser. Users log into the knowledge base, an introduction follows and subsequentlly instructions, then they can search for the diagnosis of interest (alphabetically or via first choosing the relevant organ systems), after this users receive a list of diagnoses in that system. The user chooses the appropriate diagnosis, and the device presents the elements required for a report to be considered complete. If the user clicks on an individual component, a description of how to determine, the value of that element is presented.

As it is based on HTML, ReportSupport could be provided in desktops as well. Indeed, since 2007 ReportSupport is available on the network as well. However, the investigators preferred storage of the application within the PDA because desktop computers were not always available in slide examination places and PDAs mobile connectivity was not always feasible that time. Nowadays this is not anymore an issue. The evaluation of the system, was based on an experiment involving two groups or residents; one group used the ReportSupport system, while the second did not. The performance of the first group on results reporting was evaluated three times: T0: before the availability of ReportSupport, T1: when the users had access to ReportSupport, T2: after T1 but the users did not had access to ReportSupport. The results indicated that use of the knowledgebase does not ensure report completeness, however, it was associated with complete reports and more accurate judgments of report completeness. Additionally, this performance persisted in the absence of the knowledgebase. Concluding the use of the mobile devices prooved to be a usefull tool for user training and guided towards better reporting.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

As the massive volume of cytopathology reports is related to cervical cancer prevention (via Pap Test), it is worth looking at the use of the mobile device from a gynecologist. A very recent study (S. J. Martin et al., 2015), had the purpose to assess the baseline knowledge held by training and practicing gynecologists and obstetricians on mobile applications in their field of practice. The researchers developed an electronic questionnaire to be filled by residents at selected academic institutions. According to the responses analysis, more than 80% of practitioners were using mobile phone applications in their practice. The primary use was "physician and student reference." The predominant applications used were those for cervical cytology guidelines and obstetric wheels. According to the responses, more than 80% of health care providers would like to have more applications available for use. This study proved that the use of mobile applications for gynaecologists is an expanding technology. Health care providers may choose to implement these applications into clinical practice, however it is more important to promote these applications in such a way that physicians have increased awareness. Finally the authors reported that these new technologies may be a benefit and advancement in gynecology and obstetrics both for health services provisioning as well as for medical education.

Diagnosis / Telepathology / Telecytology on the Mobile Device Using Static Images

Digital imaging in cytopathology has undergone a period of exponential growth and expansion; this was catalyzed by changes in imaging hardware and gains in computational processing. Telecytology (Della Mea et al., 2000; Markidou et al., 1999; Williams et al., 2001) is the most obvious application having the potential to be transferred directly into the mobile device. In the past, telecytology has been used for reproducibility assessment (Archondakis, 2013, 2014; Archondakis et al., 2009; Tsilalis et al., 2012), in the field of remote diagnosis there is a plethora of applications as well (Breslauer et al., 2009; Briscoe et al., 2000; Jialdasani et al., 2006; Pantanowitz et al., 2009; Pinco et al., 2009; Raab et al., 1996; Ribu et al., 2013; Spat et al., 2013; Yamashiro et al., 2004), however in all cases are used fixed computers. Nowadays there are new advances in mobile devices: first on transmission speed and second on display resolution, only these two advancements seem to allow mobile cytopathology. Applications can be extremely simple and sometimes not even requiring development of special applications, for example images may be transmitted via e-mail or uploaded on a secure web place, additionally the image viewers are already embedded into the mobile devices. Definitely there are benefits for the health sector as mobile telecytology can be a valuable tool for medical doctors, in order to manage and promote interlaboratory collaboration mainly because it allows better cytological via instant data assessment and sharing.

In the similar field of histopathology, in 2009, two independent groups (Bellina et al., 2009; McLean et al., 2009) reported on the usage of mobile phone cameras, these were used to take static digital microscopy images through the objective lens of a microscope, subsequently those images were used for telepathology. In a more specialized field, namely teledermatopathology there are numerous applications (Borve et al., 2013; Chao et al., 2013; Janda et al., 2014; Manahan et al., 2015; Massone et al., 2009; Massone et al., 2010; Wu et al., 2015), these are based on the usage of mobile phones as tools to take static dermatoscopic images and to send them for teledermatopathologic consultation. In the field of cytopathology, the image capture using the mobile phone camera, directly from the microscope is obviously a very important application.

A telemedical application is a valuable tool for cytopathologists to manage and promote interlaboratory collaboration. The result is better cytological data management and sharing, this process, in gen-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

eral, can be both user-friendly and secure. Users should be able to separate easily shared and private images, and revoke access at any time. An mHealth based application for cytopathology will be able to invite collaborators and share digital slides via email or Instant Messaging (IM). With just a few clicks, they can even embed an image in a web page for viewing. Just as easily, live users can put limitations on shared files and collaborators. Several users can review and annotate images simultaneously with digital markers, making notes immediately visible to all collaborators, and all these can be available on the mobile device.

Whole Slide Imaging using the Mobile Device

Today, digitization of entire glass slides at near the optical resolution limits of light can occur in a few minutes; additionally whole slides can be scanned in fluorescence or by multispectral imaging systems. Whole Slide Imaging (WSI) or Virtual Slides have been successfully used in surgical pathology, but its usefulness and clinical application have been limited in cytology for several reasons, mainly the lack of availability of z-axis depth focusing in cytological samples in contrast to histological have a 3D structure. However, nowadays there are available systems capable for whole slide imaging with z-axis control. This has boosted the application of digital slides for cytology (Al-Janabi et al., 2012; Fung et al., 2012; Ghaznavi et al., 2013; Gutman et al., 2013; Hipp et al., 2011; Krishnamurthy et al., 2013; Rudnisky et al., 2007; Taylor, 2011; Wright et al., 2013).

A PubMed search for whole slide imaging or digital slides reveals that there are about 200 references. However, there is no single publication that is relevant directly to cytopathology along with the use of mobile devices. The technology seems ready to support such applications and perhaps it is only a matter of timeand recruitment of appropriate resources to implement such applications.

Nowadays, there are map/navigation applications for the mobile device. For example, Google Maps, and Microsoft Bing Maps. Those applications, functionally behave extremely similar to WSI. For instance, there are used pyramidal images broken up into tiles that are served to the viewer in real-time. There are efforts using the publically available map Application Programming Interface (API) from Google and Microsoft to create WSI viewers for the desktop (Triola et al., 2011). It is obvious that this can be transferred to the mobile device as it has already happened for map/navigation. The second aspect of WSI images is depth or z-axis. This feature is required because in cytopathology doctors focus the microscope to different depths, and as mentioned earlier modern WSI systems produce multilayer scans. The capability of APIs, to present different information layers, seems compatible with this requirement. Finally, the multi-touch user interface of the mobile device screen is an additional advantage capable to provide the reach user experience.

However, WSI can be considered as just a concept. The real applications are related to the purpose that clinical doctors use WSI. Reasons for WSI include: just for archiving, for diagnosis and consultation, for providing the examination media to the patient, or just for training. Actually, there is available an application for the iPad device, this is related to the online distribution of WSIs, as teaching sets especially for countries with low resources (Fontelo et al., 2012). The study was performed via a setting that involved iPad tablet devices and fifty 3rd and 4th year medical students. There were used two web servers, providing digital pathology virtual slides via a web interface. One server was remote while the second mirrored the content on the local network. The results indicated that the speed of serving the WSIs via the local server was much faster, thus, it was preferred by the students. This study revealed the critical role of infrastructures in the acceptance of mHealth applications.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Laboratory Management Using Mobile Devices

Laboratory management operations require accurate and within time-frame transmission of critical laboratory results to the caregiver; in order to timely provide intervention and to prevent an adverse outcome. These operations are strongly linked to the provisioning of quality health services (Pouliakis et al., 2014b; Pouliakis et al., 2014c; Shen et al., 2001). Part of this process is risk management (Sciacovelli et al., 2007), either during the pre analytical phase (Vacata et al., 2007; Westbrook et al., 2008) or during sample analysis or during the post analytical phase. For this purpose, laboratories install computerized systems (Laboratory Information Systems-LIS) (Pearlman et al., 2001; Pouliakis et al., 2014b; Pouliakis et al., 2014c) to facilitate work flow and automate required actions. Mobile devices, are ideal for such applications as they are always connected and near the user. In a published study (Saw et al., 2011), the research team, reported on their experience with a setting that used SMS in reporting critical lab values. They used SMSs to notify critical laboratory results in a large teaching hospital; this application was not only to provide appropriate services; additionally it was employed to meet the documentation and audit requirements of critical result reporting, as such requirements are posed by regulatory agencies and/or ISO 15189 requirements. There was used the text messaging system (Critical Reportable Result Health care Messaging System [CRR-HMS]) that allowed a receiver to acknowledge or reject a key result by replying to a short message. When there was no response of a confirmatory receipt within 10 minutes, there were activated procedures of escalation to an alternative physician, according to a constructed roster. The benefit of this approach was that the required time for physician response was decreased from 7.3 minutes to 2 minutes. The CRR-HMS proved to be a useful tool to communicate rapidly critical results from the laboratory to targeted physicians. Eventually this enabled rapid and timely information transmission and therapeutic or patient management interventions.

Patient Doctor Interaction via the Mobile Phone or Tablet

Another aspect of mobile cytopathology is related with patient communication. Today, it is impossible to imagine people without at least one mobile device; actually in most cases this is always connected (especially in the western world as well as in developed and developing countries with high speed networks either 3G/4G or local installations of Wi-Fi). This device serves as the terminal point that patients can be informed for the availability of the cytological result (especially for the popular Papanicolaou test) or receive a reminder for a forthcoming appointment related to cytological sampling and subsequent examination. Currently, IT systems can be relatively easily interfaced and transmit automatically SMSs, e-mails or any other type notification towards the patients. For example, immediately after the cytological examination result is released from the laboratory.

Other types of interactions, between doctor and patient, are related to consulting; quite often patients have questions about their test results, especially when they are abnormal. Despite, cytopathology is a specialty that usually provides services to other medical specialties; on the other hand, in clinical cytopathology, there are patients treated directly by cytopathologists. As the examination outcomes eventually arrive at the patients' hands, they have contact information of the laboratory, if they wish to communicate with cytopathologists to be informed of the meaning of their examination outcomes from the source, rather than consulting their clinical doctor (for example the gynecologist). The mobile device is always near the patient and the physician as well. Thus it provides the media and the communication channel to interact (speak, view or message) either in real time or asynchronously.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Doctor to Doctor Interaction via the Mobile Device

In a similar manner to patient-doctor interaction; the clinical doctors collaborating with a cytopathology laboratory may receive informative messages on their mobile devices. These messages are related to the availability of a cytological examination result, requests of the laboratory for additional data that may be related to a patient under examination. Thus, mobile technology is now an open channel between cytopathology laboratories and clinical doctors.

The typical cycle of the cytopathology laboratory, starts with the sample entrance to the laboratory, along with a referral form; and ends with the production and release of the cytological report (Pouliakis et al., 2014b). Relatively old fashioned cytopathology laboratories, produce reports printed on paper, and more advanced laboratories produce electronic reports that are distributed either via sending them using secure e-mail applications or by inviting the requesting physician to login into their account and retrieve them from a safe place (usually web based). The smartphone, embeds all the capabilities to receive SMSs or emails, so that the clinical doctor is immediately informed, in order to retrieve report information.

Similarly to the patient – doctor interactions mentioned in the previous section, clinicians have numerous options to consult cytopathologists for explanatory details on the results. The most popular applications used by gynecologists are related to the interpretation of Papapnicolaou test results; thus those interactions between gynecologists and cytopathologists seem to be of the most interest. Of course, it is not only gynecologists that request services from the modern cytopathology laboratory, the vast majority of available medical specialties requests examinations from the cytopathology laboratory as well, therefore the groups of users are very large.

Use of the Mobile Device for Interaction of Personnel within the Laboratory Environment

The third aspect related to mobile communications; is the exchange of information among the cytopathology laboratory personnel. Such communication involves the doctors' requests for ancillary examinations, urgent examinations or requests to repeat an examination. Nowadays those requests are either verbal/written or in the case that the laboratory is equipped by a LIS, they are available via fixed workstations. Mobile cytopathology seems to be an attractive technology having the potential to streamline everyday routine. Laboratory personnel can be independent from computers located in fixed physical locations. In contrast, a single mobile device adapted to their preferences, a personal mobile phone, or a mobile phone provided by the laboratory, can serve this purpose. The main reason is that users do not favor to carry on and take care of more than one devices. Additionally modern devices capable to have more than one communication channels (i.e. SIM cards) can serve both private and job related tasks. The mobile device is now configurable, thus can be optimized for user personal needs/tasks, and preferences and is available on every different work location. mHealth technology, when integrated into the daily workflow, can provide exceptional consultation opportunities to distant laboratories. mHealth technology innovations can improve the professional skills of the participating medical staff and make them feel more confident in their daily work.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The Mobile Device for Quality Control and Assurance

ISO 15189:2012 for medical laboratories requires successful participation in proficiency testing programs. According to the requirements of this standard, one of the greatest challenges facing cytopathology laboratories today is the design and implementation of a board certified external quality assessment (EQA) program (proficiency testing). The purpose of the adopted program aims to ensure that microscopic (cytological) findings are correctly identified and interpreted by laboratory personnel, additionally they are stored and communicated properly (Archondakis, 2013, 2014; Friedman et al., 2006; Lee et al., 2003; Nagy et al., 2006; Vooijs et al., 1998). Nowadays EQA schemes are available as telecytological applications, for example virtual slides are stored on web sites, cytopathologists of the laboratory under certification, use their computers to remotely diagnose these virtual slides. To our opinion those applications can be easily transferred on mobile devices, thus facilitating immediate quality control and collaboration. More details on the use of mobile devices for quality control in cytopathology laboratories can be found in chapter entitled: "The Use of Mobile Health Applications for Quality Control and Accreditation Purposes in a Cytopathology Laboratory." within this book.

ISSUES, CONTROVERSIES AND PROBLEMS

As in all cases of emerging technologies, the early adopters encounter not foreseen problems. In the field of mobile cytopathology there may be numerous issues and problems.

When considering the use of WSI on a smartphone (Park et al., 2012); a major issue is that smartphone screens are small, which means physically small, because nowadays resolution of the mobile phone competes and sometimes is higher than desktop screens. The small physical size, is a requirement of the role of the device as a telephone, and thus it is not likely that future smartphones will have large enough displays (if they do then they will be no more telephones but rather tablets). A physically large display, however, is valuable and probably a requirement for viewing and interpreting a whole slide image. Screen size is not the sole issue, a second problem related to WSI and smartphones is the vendor specificity file/data formats, and these impede universal viewing in handheld devices and desktops as well. The slide scanner manufacturers use different and proprietary file formats for their whole slide images, mainly because the software that can interpret this proprietary format is the line of business for each manufacturer and a common standard is not agreed. Like in other business types cooperation among vendors is required for the benefit of the customer. Additionally, the fragmentation induced by all these proprietary formats, complicates any effort to create universal software for remote viewing of whole-slide images, especially from smartphones. Today, there are various efforts to create pathology-centric image viewers for handhelds, however, these are vendor specific (Leica Microsystems, 2011). The third issue related to WSI is related to the virtual slide size, this is enormous, ranging from more than tenths and thousands of mega bytes and up to Giga bytes, depending on the magnification, the scanned slide area and z-axis information. The large file size of WSIs seems to pose requirements for fast communication channels and device memory.

When dealing with applications on the mobile device there are numerous issues that may arise: protected patient data is stored remotely and is sent and received through the air. This makes the data vulnerable to security issues such as computer virus threats and hacker attacks. Additionally mHealth is expected to enhance usage of services related to cytopathology laboratories, thus more and more data

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

are expected to be produced. This large volume of data, especially when considering WSI, requires large storage, secured, expandable and safeguarded from failures. Similarly servers and networks require procedures to handle not only increased the load but failures as well.

SOLUTIONS AND RECOMMENDATIONS

Considering WSI on the mobile device, three issues have been identified: screen size (not resolution), proprietary formats and large file sizes. One solution to screen size could be the creation of smartphones that the screen can be expanded if needed, i.e. flexible screens. Actually there is already available a prototype by Sony announced in 2010 (SONY, 2010). This screen is so thin and flexible that can be rolled around a pencil. If smartphones with such types of screens may become widely available or the screen can be interfaced as a gadget, then, they would solve the problem of small screen size for WSI. A second solution could be the use of projectors with very small size, and thus the digital slide may be presented on the wall. Alternatively, a solution may be provided by the embedded connectors available on modern smart phones, specifically miniature HDMI connectors, that can be used to leverage a nearby large screen. As large screens are nowadays a consumer item, and available in large volumes in relatively low prices, they are available in many places, it seems that is a matter of development of the workflow and software that allows use of nearby displays from smartphones, even wirelessly.

The second issue related to WSI and smartphones; is the multitude of incompatible WSI image formats. Nowdays there is available a vendor neutral open source software library for WSI called OpenSlide (Goode A, 2008); this software library is available for desktop computers, but, porting to the mobile platform does not seem to be very difficult. Because OpenSlide is designed and implemented as a device driver, which can be found in all operating systems; it exploits vendor specific code and device resources of any type (at least this is the way expected to operate). OpenSlide supports remote viewing over the Internet as well and because the implementation is based on bindings with the OpenSeaDragon viewer (OpenSlide, 2015b) it is expected to be operational on the mobile platform as well, note that nowadays, whole slide images can be viewed over a 3G network on the smartphone browser at the OpenSlide demo site (OpenSlide, 2015a).

The third issue of WSI and the mobile device was related to the enormous image sizes. In terms of storage space, the cloud is the solution, as can provide an affordable storage space, easily expandable and moreover offers data storage assurance. Additionally no local maintenance by laboratory technicians is required. Transfer of such data volumes from/to the server towards the mobile device is an issue. However, the revolution in telecommunication networks may provide the required bandwidth. Alternatively, it is possible to transfer to the mobile device only the virtual slide parts related to the viewing fields, and the required resolution, the concept of tiles and pyramidal image formats along with multiple layers, as is the solution to reduce the volume of data required to be transmitted, ultimately viewing WSI on the mobile device may be similar to navigation application using satellite photographs, as it is already available nowadays.

In relation to the data and security issues raised by exposing data in the wireless networks as well as by exposing databases and servers in the edge of the networks, solutions may be proposed. Cloud computing, definitely is the choice (Pouliakis et al., 2014a; Pouliakis et al., 2014d) . Solutions are summarized as follows:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Threats of data security and privacy: may be anticipated by data encryption during storage and transfer and connecting with the server via encryption protocols, or via more enhanced methods proposed for the cloud environment, such as multi-cloud approach with a key sharing mechanism (Mouli et al., 2013) or patient identification cross reference numbers (Kondoh et al., 2013).
- Unauthorized access: may be anticipated by passwords and password control mechanisms and via mandatory biometric checks; note than nowadays mobile devices have the capacity for biometric security such as fingerprint recognition.
- Database safety and long-term archival: can be secured by passing this responsibility to Cloud computing service providers, in more detail:
    - Server failures may be avoided by maintaining mirror servers and/or load/ balncing
    - Efficiency of service, related to broadband speed may be ensured by multiple communication lines connected with multiple Internet service providers, thus preventing disruption of service can be assured.
    - Increased load can be easily minimized by adding more processing power to existing virtual servers and/or by splitting the application to run on multiple servers, either via load balancing or by the exploitation of multi-tier architectures and splitting to one or more separate application servers and database servers.

CONCLUSION

We have reached a point where both the software and the hardware of the mobile device platforms have matured and are powerful to be leveraged in cytopathology. The main advantage of implementing mHealth technology is that it gives all patients equal access to medical services, despite the fact that certain regions are more remote than others. This is especially important when it comes to specialists and other services usually inaccessible outside of large urban centers. mHealth technology also allows immediate collaboration among medical professionals.

Perhaps the most elegant and technically challenging application is mobile WSI. However, this does not seem to be the most used application, more demand appears to be on simple services, such as notifications, result transmission, education and perhaps laboratory management and the interactions.

One of the major benefits of the mobile platform is that provides capabilities irrelevant of time and place. The first (*freedom from time*) is delivered via the use of asynchronous communication channels: e-mail, SMSs, voice and even video mail. Simultaneously it is always possible to have a synchronous communication. The second (*freedom from place*) is related to the fact that the mobile device is personal, always on, always connected and available in every place. Therefore, there is no need for people involved in the cytopathology ecosystem to be at specific places to perform any related task.

Despite there are some security and privacy risks, that modern technology seems to have already resolved via the help of cloud computing, mHealth technology has some major benefits that the public sector and government IT organizations are would like to take advantage of.

In a very brief summary, they are as follows:

- Reduced cost (mainly for the devices)
- Flexibility: mHealth technology offers much more flexibility than past computing methods

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Mobility, employees can access information wherever they are, rather than having to remain at their desks

Work on handheld and mobile computing in cytopathology is scarce today. However, there is a great potential for new applications. Mobile hardware has nowadays become increasingly ubiquitous, reliable and fast via new devices with extremely high resolution, multicore processors, gigabytes of CPU memory and storage and multiple high-speed connectivity alternatives. Software support on such devices is comparable to that of desktop computing. However, handheld technology creates opportunities as well as challenges. Nowadays there are clear niches as well as several experimental successes with handhelds in cytopathology (especially for education purposes, telecytology, and care delivery via various communication interactions); however more study and development of standards for practicing and guidelines for validation are required. In addition, there are still questions related to security. Even so, it is likely that mobile computing will play a large role in cytopathology in the forthcoming digital decade. mHealth seems to have the potential to change the picture of future cytopathology and pathology.

REFERENCES

Al-Janabi, S., Huisman, A., Willems, S. M., & Van Diest, P. J. (2012). Digital slide images for primary diagnostics in breast pathology: A feasibility study. *Human Pathology*, *43*(12), 2318–2325. doi:10.1016/j.humpath.2012.03.027 PMID:22901465

Archondakis, S. (2013). The Use of Static Telemedical Applications of Cytopathology for Proficiency Testing. *International Journal of Reliable and Quality E-Healthcare*, *2*(2), 47–53. doi:10.4018/ijrqeh.2013040104

Archondakis, S. (2014). Static Telecytological Applications for Proficiency Testing. In A. Information Resources Management (Ed.), Nanotechnology: Concepts, Methodologies, Tools, and Applications (pp. 556-568). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-5125-8.ch023

Archondakis, S., Georgoulakis, J., Stamataki, M., Anninos, D., Skagias, L., Panayiotides, I., ... Karakitsos, P. (2009). Telecytology: A tool for quality assessment and improvement in the evaluation of thyroid fine-needle aspiration specimens. *Telemedicine Journal and e-Health*, *15*(7), 713–717. doi:10.1089/tmj.2009.0037 PMID:19694595

Banitsas, K., Perakis, K., Koutsouris, D., Konis, G., & Tachakra, S. (2005). Ambulance 3G. *Conference Proceedings; ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *1*, 371–374. PMID:17282191

Banitsas, K. A., Perakis, K., Tachakra, S., & Koutsouris, D. (2006). Use of 3G mobile phone links for teleconsultation between a moving ambulance and a hospital base station. *Journal of Telemedicine and Telecare*, *12*(1), 23–26. doi:10.1258/135763306775321335 PMID:16438775

Bellina, L., & Missoni, E. (2009). Mobile cell-phones (M-phones) in telemicroscopy: Increasing connectivity of isolated laboratories. *Diagnostic Pathology*, *4*(1), 19. doi:10.1186/1746-1596-4-19 PMID:19545373

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Borve, A., Terstappen, K., Sandberg, C., & Paoli, J. (2013). Mobile teledermoscopy-there's an app for that! *Dermatol Pract Concept*, *3*(2), 41–48. doi:10.5826/dpc.0302a05 PMID:23785643

Breslauer, D. N., Maamari, R. N., Switz, N. A., Lam, W. A., & Fletcher, D. A. (2009). Mobile phone based clinical microscopy for global health applications. *PLoS ONE*, *4*(7), e6320. doi:10.1371/journal.pone.0006320 PMID:19623251

Briscoe, D., Adair, C. F., Thompson, L. D., Tellado, M. V., Buckner, S. B., Rosenthal, D. L., & O'Leary, T. J. (2000). Telecytologic diagnosis of breast fine needle aspiration biopsies. Intraobserver concordance. *Acta Cytologica*, *44*(2), 175–180. doi:10.1159/000326357 PMID:10740603

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research*, *13*(3), e55. doi:10.2196/jmir.1838 PMID:21840837

Chao, J. T. II, Loescher, L. J., Soyer, H. P., & Curiel-Lewandrowski, C. (2013). Barriers to mobile teledermoscopy in primary care. *Journal of the American Academy of Dermatology*, *69*(5), 821–824. doi:10.1016/j.jaad.2012.10.039 PMID:24124819

Choudhri, A. F., Carr, T. M. III, Ho, C. P., Stone, J. R., Gay, S. B., & Lambert, D. L. (2012). Handheld device review of abdominal CT for the evaluation of acute appendicitis. *Journal of Digital Imaging*, *25*(4), 492–496. doi:10.100710278-011-9431-9 PMID:22146833

Choudhri, A. F., Norton, P. T., Carr, T. M. III, Stone, J. R., Hagspiel, K. D., & Dake, M. D. (2013). Diagnosis and treatment planning of acute aortic emergencies using a handheld DICOM viewer. *Emergency Radiology*, *20*(4), 267–272. doi:10.100710140-013-1118-8 PMID:23525945

Clarke, J., Proudfoot, J., Birch, M. R., Whitton, A. E., Parker, G., Manicavasagar, V., ... Hadzi-Pavlovic, D. (2014). Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: Secondary analysis of a randomised controlled trial. *BMC Psychiatry*, *14*(1), 272. doi:10.118612888-014-0272-1 PMID:25252853

Collins, B. (2011). iPad and cytopathology: Content creation and delivery. *CytoJournal*, *8*(16c). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3230889/

Crowe, A., Linder, A., Hameed, O., Salih, C., Roberson, J., Gidley, J., & Eltoum, I. A. (2011). The impact of implementation of the Bethesda System for Reporting Thyroid Cytopathology on the quality of reporting, "risk" of malignancy, surgical rate, and rate of frozen sections requested for thyroid lesions. *Cancer Cytopathol*, *119*(5), 315–321. doi:10.1002/cncy.20174 PMID:21751427

Della Mea, V., Cataldi, P., Pertoldi, B., & Beltrami, C. A. (2000). Combining dynamic and static robotic telepathology: A report on 184 consecutive cases of frozen sections, histology and cytology. *Analytical Cellular Pathology*, *20*(1), 33–39. doi:10.1155/2000/734943 PMID:11007436

Elsevier. (2015). *Robbins Pathology Flash Cards.* Retrieved 15 April, 2015, from https://itunes.apple.com/us/app/robbins-pathology-flash-cards/id367031323?mt=8

Finkelstein, J., Hripcsak, G., & Cabrera, M. (1998). Telematic system for monitoring of asthma severity in patients' homes. *Studies in Health Technology and Informatics*, *52*(Pt 1), 272–276. PMID:10384460

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fontelo, P., Faustorilla, J., Gavino, A., & Marcelo, A. (2012). Digital pathology - implementation challenges in low-resource countries. *Anal Cell Pathol (Amst)*, *35*(1), 31–36. doi:10.1155/2012/125695 PMID:22233702

Friedman, C. P., & Wyatt, J. (2006). *Evaluation methods in medical informatics* (2nd ed.). New York: Springer.

Fung, K. M., Hassell, L. A., Talbert, M. L., Wiechmann, A. F., Chaser, B. E., & Ramey, J. (2012). Whole slide images and digital media in pathology education, testing, and practice: The Oklahoma experience. *Analytical cellular pathology (Amsterdam)*, *35*(1), 37–40. doi:10.1155/2012/103857 PMID:21965282

Ghaznavi, F., Evans, A., Madabhushi, A., & Feldman, M. (2013). Digital imaging in pathology: Whole-slide imaging and beyond. *Annual Review of Pathology*, *8*(1), 331–359. doi:10.1146/annurev-pathol-011811-120902 PMID:23157334

Ghorai, K., Akter, S., Khatun, F., & Ray, P. (2014). mHealth for Smoking Cessation Programs: A Systematic Review. *J Pers Med*, *4*(3), 412–423. doi:10.3390/jpm4030412 PMID:25563359

Goode, A. S. M. (2008). *A vendor-neutral library and viewer for wholeslide images* (Vol. Tech Rep CMU-CS-08-136.). Pittsburgh PA: School of Computer Science, Carnegie Mellon University.

Gupta, S., Chang, P., Anyigbo, N., & Sabharwal, A. (2011). *mobileSpiro: accurate mobile spirometry for self-management of asthma*. Paper presented at the First ACM Workshop on Mobile Systems, Applications, and Services for Healthcare, Seattle, WA. 10.1145/2064942.2064944

Gutman, D. A., Cobb, J., Somanna, D., Park, Y., Wang, F., Kurc, T., ... Cooper, L. A. (2013). Cancer Digital Slide Archive: An informatics resource to support integrated in silico analysis of TCGA pathology data. *Journal of the American Medical Informatics Association*, *20*(6), 1091–1098. doi:10.1136/amiajnl-2012-001469 PMID:23893318

Hipp, J. D., Lucas, D. R., Emmert-Buck, M. R., Compton, C. C., & Balis, U. J. (2011). Digital slide repositories for publications: Lessons learned from the microarray community. *The American Journal of Surgical Pathology*, *35*(6), 783–786. doi:10.1097/PAS.0b013e31821946b6 PMID:21552111

Hirsch, D., Robenshtok, E., Bachar, G., Braslavsky, D., & Benbassat, C. (2015). The Implementation of the Bethesda System for Reporting Thyroid Cytopathology Improves Malignancy Detection Despite Lower Rate of Thyroidectomy in Indeterminate Nodules. *World Journal of Surgery*, *39*(8), 1959–1965. doi:10.100700268-015-3032-6 PMID:25809059

Janda, M., Loescher, L. J., Banan, P., Horsham, C., & Soyer, H. P. (2014). Lesion selection by melanoma high-risk consumers during skin self-examination using mobile teledermoscopy. *JAMA Dermatology*, *150*(6), 656–658. doi:10.1001/jamadermatol.2013.7743 PMID:24522284

Jialdasani, R., Desai, S., Gupta, M., Kothari, A., Deshpande, R., Shet, T., ... Chinoy, R. (2006). An analysis of 46 static telecytology cases over a period of two years. *Journal of Telemedicine and Telecare*, *12*(6), 311–314. doi:10.1258/135763306778558132 PMID:17022840

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Johnson, P. T., Zimmerman, S. L., Heath, D., Eng, J., Horton, K. M., Scott, W. W., & Fishman, E. K. (2012). The iPad as a mobile device for CT display and interpretation: Diagnostic accuracy for identification of pulmonary embolism. *Emergency Radiology*, *19*(4), 323–327. doi:10.100710140-012-1037-0 PMID:22450843

Kiselev, A. R., Gridnev, V. I., Shvartz, V. A., Posnenkova, O. M., & Dovgalevsky, P. Y. (2012). Active ambulatory care management supported by short message services and mobile phone technology in patients with arterial hypertension. *Journal of the American Society of Hypertension*, *6*(5), 346–355. doi:10.1016/j.jash.2012.08.001 PMID:22995803

Kondoh, H., Teramoto, K., Kawai, T., Mochida, M., & Nishimura, M. (2013). Development of the Regional EPR and PACS Sharing System on the Infrastructure of Cloud Computing Technology Controlled by Patient Identifier Cross Reference Manager. *Studies in Health Technology and Informatics*, *192*, 1073. PMID:23920847

Krishnamurthy, S., Mathews, K., McClure, S., Murray, M., Gilcrease, M., Albarracin, C., ... Liang, K. (2013). Multi-Institutional Comparison of Whole Slide Digital Imaging and Optical Microscopy for Interpretation of Hematoxylin-Eosin-Stained Breast Tissue Sections. *Archives of Pathology & Laboratory Medicine*, *137*(12), 1733–1739. doi:10.5858/arpa.2012-0437-OA PMID:23947655

Lee, E. S., Kim, I. S., Choi, J. S., Yeom, B. W., Kim, H. K., Han, J. H., ... Leong, A. S. (2003). Accuracy and reproducibility of telecytology diagnosis of cervical smears. A tool for quality assurance programs. *American Journal of Clinical Pathology*, *119*(3), 356–360. doi:10.1309/7YTVAG4XNR48T75H PMID:12645336

Leica Microsystems. (2011). *Leica Microsystems Launches Free Digital Pathology App for iPad and iPhone*. Retrieved 10 April, 2015, from http://www.leica-microsystems.com/news-media/news/news-details/article/leica-microsystems-launches-free-digital-pathology-app-for-ipad-and-iphone/

Manahan, M. N., Soyer, H. P., Loescher, L. J., Horsham, C., Vagenas, D., Whiteman, D. C., ... Janda, M. (2015). A pilot trial of mobile, patient-performed teledermoscopy. *The British Journal of Dermatology*, *172*(4), 1072–1080. doi:10.1111/bjd.13550 PMID:25418126

Markidou, S., Karakitsos, P., & Pouliakis, A. (1999). Telecytology. In O. Ferrer-Roca & M. Sosa-Iudicissa (Eds.), *Handbook of Telemedicine* (Vol. 54). Amsterdam: IOS press.

Martin, S. J., Chen, G., Dattilo, J., & Johnson, C. T. (2015). Physicians' Knowledge Regarding Tablet- and Handheld-Based Applications in Gynecology and Obstetrics. *Obstetrics and Gynecology*, *125*(Suppl 1), 119S–120S. doi:10.1097/01.AOG.0000463206.49148.98

Martin, T. (2012). Assessing mHealth: Opportunities and barriers to patient engagement. *Journal of Health Care for the Poor and Underserved*, *23*(3), 935–941. doi:10.1353/hpu.2012.0087 PMID:24212144

Massone, C., Brunasso, A. M., Campbell, T. M., & Soyer, H. P. (2009). Mobile teledermoscopy--melanoma diagnosis by one click? *Seminars in Cutaneous Medicine and Surgery*, *28*(3), 203–205. doi:10.1016/j.sder.2009.06.002 PMID:19782945

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Massone, C., Brunasso, A. M., Hofmann-Wellenhof, R., Gulia, A., & Soyer, H. P. (2010). Teledermoscopy: Education, discussion forums, teleconsulting and mobile teledermoscopy. *Giornale Italiano di Dermatologia e Venereologia*, *145*(1), 127–132. PMID:20197751

McLean, R., Jury, C., Bazeos, A., & Lewis, S. M. (2009). Application of camera phones in telehaematology. *Journal of Telemedicine and Telecare*, *15*(7), 339–343. doi:10.1258/jtt.2009.090114 PMID:19815902

Meszaros, L. (2014). *Johns Hopkins Atlas of Pancreatic Cytopathology*. Retrieved 10 February, 2015, from http://www.mdlinx.com/endocrinology/mobile-app-center/mobile-app-detail.cfm?app_id=1221

Mouli, K. C., & Sesadri, U. (2013). Single to Multi Clouds for Security in Cloud Computing by using Secret Key Sharing. *Internatioonal Journal of Computers and Technology*, *10*(4), 1539–1545.

Mun, S. K., Elsayed, A. M., Tohme, W. G., & Wu, Y. C. (1995). Teleradiology/telepathology requirements and implementation. *Journal of Medical Systems*, *19*(2), 153–164. doi:10.1007/BF02257066 PMID:7602247

Nagy, G. K., & Newton, L. E. (2006). Cytopathology proficiency testing: Where do we go from here? *Diagnostic Cytopathology*, *34*(4), 257–264. doi:10.1002/dc.20361 PMID:16544329

Ng, C. H., & Yeo, J. F. (2004). The role of Internet and personal digital assistant in oral and maxillofacial pathology. *Annals of the Academy of Medicine, Singapore*, *33*(4Suppl), 50–52. PMID:15389308

OpenSlide. (2015a). *OpenSlide Demo Site*. Retrieved 25 May, 2015, from http://www.openslide.org/demo

OpenSlide. (2015b). *OpenSlide homepage*. Retrieved 10 April, 2015, from http://www.openslide.org/

Pandher, P. S., & Bhullar, K. K. (2014). Smartphone applications for seizure management. *Health Informatics Journal*. doi:10.1177/1460458214540906 PMID:25038202

Pantanowitz, L., Henricks, W. H., & Beckwith, B. A. (2007). Medical laboratory informatics. *Clinics in Laboratory Medicine*, *27*(4), 823–843. doi:10.1016/j.cll.2007.07.011 PMID:17950900

Pantanowitz, L., Hornish, M., & Goulart, R. A. (2009). The impact of digital imaging in the field of cytopathology. *CytoJournal*, *6*(1), 6. doi:10.4103/1742-6413.48606 PMID:19495408

Park, S., Parwani, A., Satyanarayanan, M., & Pantanowitz, L. (2012). Handheld computing in pathology. *J Pathol Inform*, *3*(1), 15–15. doi:10.4103/2153-3539.95127 PMID:22616027

Patrick, K., Raab, F., Adams, M. A., Dillon, L., Zabinski, M., Rock, C. L., ... Norman, G. J. (2009). A text message-based intervention for weight loss: Randomized controlled trial. *Journal of Medical Internet Research*, *11*(1), e1. doi:10.2196/jmir.1100 PMID:19141433

Pavlopoulos, S., Kyriacou, E., Berler, A., Dembeyiotis, S., & Koutsouris, D. (1998). A novel emergency telemedicine system based on wireless communication technology--AMBULANCE. *IEEE Transactions on Information Technology in Biomedicine*, *2*(4), 261–267. doi:10.1109/4233.737581 PMID:10719536

Pearlman, E. S., Wolfert, M. S., Miele, R., Bilello, L., & Stauffer, J. (2001). Utilization management and information technology: Adapting to the new era. *Clinical Leadership & Management Review*, *15*(2), 85–88. PMID:11299910

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Pinco, J., Goulart, R. A., Otis, C. N., Garb, J., & Pantanowitz, L. (2009). Impact of digital image manipulation in cytology. *Archives of Pathology & Laboratory Medicine*, *133*(1), 57–61. PMID:19123737

Pouliakis, A., Archondakis, S., Karakitsou, E., & Karakitsos, P. (2014a). Cloud Computing for Cytopathologists. In A. Moumtzoglou & A. Kastania (Eds.), *Cloud computing applications for quality health care delivery* (pp. 250–271). Hershey, PA: IGI Global.

Pouliakis, A., Athanasiadi, E., Karakitsou, E., Archondakis, S., Mourtzikou, A., Stamouli, M., ... Karakitsos, P. (2014b). ISO 15189:2012 Management Requirements for Cytopathology Laboratory Information Systems. *International Journal of Reliable and Quality E-Healthcare*, *3*(3), 37–57. doi:10.4018/ijrqeh.2014070103

Pouliakis, A., Margari, N., Spathis, A., Kottaridi, C., Stamouli, M., Mourtzikou, A., ... Karakitsos, P. (2014c). ISO 15189:2012 Technical Requirements for Cytopathology Laboratory Information Systems. *International Journal of Reliable and Quality E-Healthcare*, *3*(3), 58–80. doi:10.4018/ijrqeh.2014070104

Pouliakis, A., Spathis, A., Kottaridi, C., Mourtzikou, A., Stamouli, M., Archondakis, S., ... Karakitsos, P. (2014d). Cloud Computing for BioLabs. In A. Moumtzoglou & A. Kastania (Eds.), *Cloud computing applications for quality health care delivery* (pp. 228–249). Hershey, PA: IGI Global.

Quinn, C. C., Shardell, M. D., Terrin, M. L., Barr, E. A., Ballew, S. H., & Gruber-Baldini, A. L. (2011). Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care*, *34*(9), 1934–1942. doi:10.2337/dc11-0366 PMID:21788632

Raab, S. S., Zaleski, M. S., Thomas, P. A., Niemann, T. H., Isacson, C., & Jensen, C. S. (1996). Telecytology: Diagnostic accuracy in cervical-vaginal smears. *American Journal of Clinical Pathology*, *105*(5), 599–603. PMID:8623769

Ribu, L., Holmen, H., Torbjornsen, A., Wahl, A. K., Grottland, A., Smastuen, M. C., ... Arsand, E. (2013). Low-intensity self-management intervention for persons with type 2 diabetes using a mobile phone-based diabetes diary, with and without health counseling and motivational interviewing: Protocol for a randomized controlled trial. *JMIR Res Protoc*, *2*(2), e34. doi:10.2196/resprot.2768 PMID:23978690

Rosales Saurer, B., Mueller-Gorchs, M., & Kunze, C. (2009). Scenario-based design of an ICT platform for mobile information services in ambulatory care nursing. *Studies in Health Technology and Informatics*, *146*, 64–68. PMID:19592810

Rudnisky, C. J., Tennant, M. T., Weis, E., Ting, A., Hinz, B. J., & Greve, M. D. (2007). Web-based grading of compressed stereoscopic digital photography versus standard slide film photography for the diagnosis of diabetic retinopathy. *Ophthalmology*, *114*(9), 1748–1754. doi:10.1016/j.ophtha.2006.12.010 PMID:17368543

Sama, P. R., Eapen, Z. J., Weinfurt, K. P., Shah, B. R., & Schulman, K. A. (2014). An evaluation of mobile health application tools. *JMIR Mhealth Uhealth*, *2*(2), e19. doi:10.2196/mhealth.3088 PMID:25099179

Saw, S., Loh, T. P., Ang, S. B., Yip, J. W., & Sethi, S. K. (2011). Meeting regulatory requirements by the use of cell phone text message notification with autoescalation and loop closure for reporting of critical laboratory results. *American Journal of Clinical Pathology*, *136*(1), 30–34. doi:10.1309/AJCPUZ53XZWQFYIS PMID:21685029

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sciacovelli, L., Secchiero, S., Zardo, L., D'Osualdo, A., & Plebani, M. (2007). Risk management in laboratory medicine: Quality assurance programs and professional competence. *Clinical Chemistry and Laboratory Medicine*, *45*(6), 756–765. doi:10.1515/CCLM.2007.165 PMID:17579529

Sharma, P., & Kamal, V. (2006). A novel application of mobile phone cameras in undergraduate pathology education. *Medical Teacher*, *28*(8), 745–746. PMID:17642089

Shen, Z., & Yang, Z. (2001). The problems and strategy relevant to the quality management of clinical laboratories. *Clinical Chemistry and Laboratory Medicine*, *39*(12), 1216–1218. doi:10.1515/CCLM.2001.194 PMID:11798079

Skeate, R. C., Wahi, M. M., Jessurun, J., & Connelly, D. P. (2007). Personal digital assistant-enabled report content knowledgebase results in more complete pathology reports and enhances resident learning. *Human Pathology*, *38*(12), 1727–1735. doi:10.1016/j.humpath.2007.05.019 PMID:17949779

Skrovseth, S. O., Arsand, E., Godtliebsen, F., & Hartvigsen, G. (2012). Mobile phone-based pattern recognition and data analysis for patients with type 1 diabetes. *Diabetes Technology & Therapeutics*, *14*(12), 1098–1104. doi:10.1089/dia.2012.0160 PMID:23035775

SONY. (2010). *Sony Develops a "Rollable" OTFT*1-driven OLED Display that can wrap around a Pencil*. Retrieved 24 May, 2015, from http://www.sony.net/SonyInfo/News/Press/201005/10-070E/

Spat, S., Holl, B., Petritsch, G., Schaupp, L., Beck, P., & Pieber, T. R. (2013). Automatic system testing of a decision support system for insulin dosing using Google Android. *Studies in Health Technology and Informatics*, *186*, 187–191. PMID:23542995

Stergiou, N., Georgoulakis, G., Margari, N., Aninos, D., Stamataki, M., Stergiou, E., ... Karakitsos, P. (2009). Using a web-based system for the continuous distance education in cytopathology. *International Journal of Medical Informatics*, *78*(12), 827–838. doi:10.1016/j.ijmedinf.2009.08.007 PMID:19775933

Taylor, C. R. (2011). From microscopy to whole slide digital images: A century and a half of image analysis. *Applied Immunohistochemistry & Molecular Morphology*, *19*(6), 491–493. doi:10.1097/PAI.0b013e318229ffd6 PMID:22089486

Toomey, R. J., Ryan, J. T., McEntee, M. F., Evanoff, M. G., Chakraborty, D. P., McNulty, J. P., ... Brennan, P. C. (2010). Diagnostic efficacy of handheld devices for emergency radiologic consultation. *AJR. American Journal of Roentgenology*, *194*(2), 469–474. doi:10.2214/AJR.09.3418 PMID:20093611

Triola, M. M., & Holloway, W. J. (2011). Enhanced virtual microscopy for collaborative education. *BMC Medical Education*, *11*(1), 4. doi:10.1186/1472-6920-11-4 PMID:21269474

Tsilalis, T., Archondakis, S., Meristoudis, C., Margari, N., Pouliakis, A., Skagias, L., ... Karakitsos, P. (2012). Assessment of Static Telecytological Diagnoses' Reproducibility in Cervical Smears Prepared by Means of Liquid-Based Cytology. *Telemedicine Journal and e-Health*, *18*(7), 516–520. doi:10.1089/tmj.2011.0167 PMID:22856666

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Vacata, V., Jahns-Streubel, G., Baldus, M., & Wood, W. G. (2007). Practical solution for control of the pre-analytical phase in decentralized clinical laboratories for meeting the requirements of the medical laboratory accreditation standard DIN EN ISO 15189. *Clinical Laboratory*, *53*(3-4), 211–215. PMID:17447659

Vaisanen, O., Makijarvi, M., & Silfvast, T. (2003). Prehospital ECG transmission: Comparison of advanced mobile phone and facsimile devices in an urban Emergency Medical Service System. *Resuscitation*, *57*(2), 179–185. doi:10.1016/S0300-9572(03)00028-5 PMID:12745186

Vooijs, G. P., Davey, D. D., Somrak, T. M., Goodell, R. M., Grohs, D. H., Knesel, E. A. Jr, ... Wilbur, D. C. (1998). Computerized training and proficiency testing. International Academy of Cytology Task Force summary. Diagnostic Cytology Towards the 21st Century: An International Expert Conference and Tutorial. *Acta Cytologica*, *42*(1), 141–147. doi:10.1159/000331539 PMID:9479333

Weinstein, L. J., Epstein, J. I., Edlow, D., & Westra, W. H. (1997). Static image analysis of skin specimens: The application of telepathology to frozen section evaluation. *Human Pathology*, *28*(1), 30–35. doi:10.1016/S0046-8177(97)90275-6 PMID:9013828

Westbrook, J. I., Georgiou, A., & Rob, M. I. (2008). Computerised order entry systems: Sustained impact on laboratory efficiency and mortality rates? *Studies in Health Technology and Informatics*, *136*, 345–350. PMID:18487755

Williams, B. H., Mullick, F. G., Butler, D. R., Herring, R. F., & O'Leary, T. J. (2001). Clinical evaluation of an international static image-based telepathology service. *Human Pathology*, *32*(12), 1309–1317. doi:10.1053/hupa.2001.29649 PMID:11774162

Wright, A. M., Smith, D., Dhurandhar, B., Fairley, T., Scheiber-Pacht, M., Chakraborty, S., ... Coffey, D. M. (2013). Digital slide imaging in cervicovaginal cytology: A pilot study. *Archives of Pathology & Laboratory Medicine*, *137*(5), 618–624. doi:10.5858/arpa.2012-0430-OA PMID:22970841

Wu, X., Oliveria, S. A., Yagerman, S., Chen, L., DeFazio, J., Braun, R., & Marghoob, A. A. (2015). Feasibility and Efficacy of Patient-Initiated Mobile Teledermoscopy for Short-term Monitoring of Clinically Atypical Nevi. *JAMA Dermatology*, *151*(5), 489. doi:10.1001/jamadermatol.2014.3837 PMID:25629626

Yamashiro, K., Kawamura, N., Matsubayashi, S., Dota, K., Suzuki, H., Mizushima, H., ... Azumi, N. (2004). Telecytology in Hokkaido Island, Japan: Results of primary telecytodiagnosis of routine cases. *Cytopathology*, *15*(4), 221–227. doi:10.1111/j.1365-2303.2004.00147.x PMID:15324451

Ybarra, M. L., Holtrop, J. S., Prescott, T. L., & Strong, D. (2014). Process evaluation of a mHealth program: Lessons learned from Stop My Smoking USA, a text messaging-based smoking cessation program for young adults. *Patient Education and Counseling*, *97*(2), 239–243. doi:10.1016/j.pec.2014.07.009 PMID:25103183

Zerth, J., Besser, J., & Reichert, A. (2012). Effectiveness and efficiency for ambulatory care assisted by mobile technological devices. *Biomed Tech (Berl), 57*(1).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Cytopathology: A specialty of medicine related to the study and diagnosis of diseases by the examination of cells.

E-Health: Is the healthcare supported by electronics, informatics and tele-communications.

E-Learning: Is a broad concept referring to the application of information and communication technologies (ICT) for learning purposes.

Laboratory Information System (LIS): Or Laboratory Information Management System (LIMS) is a software-based system for the support of operations in a modern laboratory, such as workflow, sample tracking, data exchange interfaces. LIS systems are often capable to be connected with medical analyzers for automated extraction and storage of measurements.

Mobile Health (mHealth): The practice of medicine and public health supported by mobile devices.

Quality Control: The set of processes by which entities review the quality of all factors involved in product, service or activity. During QC processes, the products, services and activities are tested or validated in order to reveal defects and problems, before their release.

Telecytology: The application of cytopathology from distance.

Telediagnosis: Is the determination of a disease at a site remote from the patient based on transmitted data.

Virtual Slide: or Whole Slide Imaging, a virtual slide is created when a glass slide is entirely scanned digitally.

This research was previously published in M-Health Innovations for Patient-Centered Care edited by Anastasius Moumtzoglou; pages 240-261; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 97
Examining the Influence of Information Technology on Modern Health Care

Edward T. Chen
University of Massachusetts, Lowell, USA

ABSTRACT

Health care costs continue to rise at a level that far exceeds the rate of inflation. IT will be necessary in the computation and organization of complex algorithms presented in bundled payments and other initiatives. Currently in health care, a patient's medical history is not easily accessible by physicians and other medical personnel. IT can play the pivotal role in rectifying this problem in tracking the record in a universally designed environment. Advanced databases are needed to integrate facilities within health care systems. This chapter is to explore the current framework of Information Technology in the U.S. health care industry and to examine the topic covering the following areas: (a) IT's influence on the Affordable Care Act, (b) the emergence of the Electronic Health Record (EHR), also known as the Electronic Medical Record (EMR), and (c) the integration of databases across health care organizations through advanced systems like Epic.

INTRODUCTION

The road to change in the health care industry is being paved currently by Health Care Reform. More specifically, the framework of the Patient Protection and Affordable Care Act (PPACA) has held stakeholders in the health care industry accountable for implementing IT modification within their organizations, government bodies, and society as a whole. Health care has an abundance of concentrations, all of which require some form of information technology resource in order for these bodies to achieve the highest service quality of patient care (Rosenbaum, & Margulies, 2011). The purpose of this chapter paper is to explore numerous avenues in the field of health care and to examine the impact that IT has on the U.S. health care industry.

DOI: 10.4018/978-1-7998-1204-3.ch097

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The advancement of information technology in health care is pivotal to the guidelines set in the Patient Protection and Affordable Care Act. As health care organizations have become clinically data driven, the need for information systems and IT is imperative (Vo & Bhaskar, 2012). Analysts in IT departments at hospitals play the role in maintaining and implementing systems. Clinical staff members, such as nurses and physicians, play the role of end users and require the skills needed to effectively operate these systems and forms of technology in order to carry out their clinical duties (Najaftorkaman, Ghapanchi, Talaei-Khoei, & Ray, 2015). Key developments in IT include the transformation of the medical record from paper to digital, the use of mobile devices, and the emergence of integrated health care systems that have the capacity for handling the many complexities that the industry has to offer.

Equally impacted are all U.S. employers who offer health insurance to their employees and the dependents of their employees. While employers attempt to contain the trend of health insurance expense, information systems are needed to pinpoint conditions that prove to be the most costly and implement benefit redesign plans to reduce costs and promote wellness. Employers also keep their staff productive by keeping them healthy. In the same light, information systems help employees examine disease management programs that are geared towards steering employees to engage in healthy behavior and seek preventive care regularly. If health care was once described as being technically challenged when compared with other industries, that notion is being put to rest by the progression of IT in the industry (Vo & Bhaskar, 2012).

For health care to withstand the winds of change, innovation must be at the forefront of the movement. The industry has endured recent developments in IT, but new ideas and expansion in IT staff are imperative for Information Technology to continue to positively impact the health care industry. Ultimately, the goal of a health care organization should be to provide the best possible care to the patient and ensure the well-being of its surrounding community (Ku, Jones, Finnegan, Shin, & Rosenbaum, 2009).

RECENT GOVERNMENT DEVELOPMENTS ON HEALTH CARE IT

The Patient Protection and Affordable Care Act (PPACA) was signed by President Barack Obama on March 3, 2010 and upheld by Congress on June 28, 2012. The act mandates that all Americans are required to be insured with health coverage or pay a penalty to the government beginning in 2014. Another provision includes the expansion of Medicaid to adults whose incomes are at a defined rate at or below the federal poverty limit. Furthermore, the act seeks to decrease the cost of health care (Nussbaum, Tirrell, Wechsler, & Randall, 2010; Rosenbaum, & Margulies, 2011).

Exchanges have been established in the majority of the states either on their own or with the assistance of the Federal Government. The exchanges represent online marketplaces where participants can shop for health insurance plans. The exchanges also provide online resources that allow users to determine their eligibility for a federal subsidy (Carey & Appleby, 2013; Najaftorkaman, et al., 2015).

Healthcare.gov

Many organizations that offer medical insurance as an employee benefit to their staff have adopted online self-service enrollment options. Employees simply log in to an HR system, such as PeopleSoft, and enroll in coverage during their open enrollment period, as new hires, or in lieu of experiencing a qualifying event. Enrolling in coverage through an exchange works in a similar manner. A participant

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

visits Healthcare.gov or a state run exchange and then completes an online application. The system determines the eligibility for Medicaid or a subsidy from the government to offset the full price of commercial insurance coverage. The participant can then enroll in government sponsored or commercial insurance products (Ford, 2013).

Healthcare.gov was scheduled to go live on October 1, 2013 and enrollment in a plan would be effective as of January 1, 2014. However, the website failed. Visitors experienced pages that would not load and were advised through screen prompts to be patient. President Obama's administration announced that the site's issues would be resolved quickly. However, a resolution was not immediately reached and issues began to progressively grow. At this time, the site had been operating at about 43 percent. As a result of strenuous technical maintenance, the site began to operate at 90 percent but not until the beginning of December 2013 (Somashekhar & Sun, 2013).

So what happened? Experts in the field might argue that complex IT systems like the one rolled out with Healthcare.gov are prone to begin with a low level of success upon early stages of operation by end users. In a November 2013 article of *ComputerWorld* magazine, Lev Lesokhin, Senior VP at Cast Software, suggested that the complexity of Healthcare.gov could have been measured in the magnitude of the system's lines of code. Lesokhin estimated this number to be around 500,000. In the same article, author Patrick Thibodeau reports on the findings of Standish Group, a research services organization that specializes in uncovering successes and failures in IT projects. Of the 3,555 government and commercial IT projects from 2003 to 2012 in its database, Standish Group found that 52% were challenged, 41.4% were failures and only 6.4% were considered successful. Ultimately, the researchers came to the conclusion that Healthcare.gov had only a 6.4% chance of being successful. The site was destined to fail at some level (Thibodeau & Gross, 2013).

With enrollments needed to be made at the end of 2013 for a January 1, 2014 effective date, Healhcare.gov had limited time for development and testing phases. Regardless of Standish Group's findings, it is possible that developers did not have adequate time frames to complete testing phases that may have been critical in avoiding the initial failure.

IT in the Accountable Care Organization

Section 3022 of the PPACA introduces a Medicare Shared Savings Program for the Accountable Care Organization (ACO). An Accountable Care Organization is a health care provider that aims at delivering high service quality care and reducing the cost of care. In doing so, the ACO can expect to achieve an incentive in the form of shared Medicare savings (Berwick, 2011).

The ACO model depends greatly on Health Information Technology (HIT). HIT has evolved in the health care industry more and more as data becomes the centerpiece of delivering high service quality care, cost analysis and payment reform. The ACO concept links care and payment risk across multiple organizations and physicians. In order for the model to survive, data sharing will be crucial. The delivery and storage of this data is where technical resources come into play. The ACO portrays a team concept where multiple entities collaborate to achieve financial incentive while providing the best possible care. In order to do so, a patient's Electronic Health Record (EHR) must be electronically maintained and available across all clinical professionals involved. The EHR concept is at the early stages of development in health care and currently represents limited data sharing capabilities. It is evident that Accountable Care Organizations will be required to invest in IT interfaces that track the patients across the entire episode

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of care (Bitton & Filer, 2012). Clinical data is at the forefront of the ACO model. As organizations adopt the concept, they will require sophisticated IT resources to share data internally and externally with other acute and post-acute care facilities.

IT in the Patient-Centered Medical Home

The PPACA supports the establishment of the Patient Centered Medical Home (PCMH). The concept of the PCMH is that a single organization delivers comprehensive and primary care to children, youth and adults. This model intends to establish high service quality care for the patient while coordinating specialty care with supporting health care organizations in the community. The structure is such that each patient has a personal physician who leads a team of clinical staff within the PCMH. The clinical staff is responsible for providing ongoing care to the patient (U.S. Department of Health & Human Services, 2014). Ultimately, a medical home should eliminate gaps and foster care coordination that traditional primary care practices may be unable to accomplish. The PPACA supports the concept by making grants available to eligible providers that fit certain PCMH criteria (Epstein, Fiscella, Lesser, & Stange, 2010).

For medical homes to be successful, modern technology should be embraced in the patient to provider relationship. Practices benefit from data analysis and technological communication in the form of decision support systems, electronic registries, systems that facilitate the transition of care, and telehealth (Bates & Bitton, 2010; Fleurant, Kell, Love, Jenter, Volk, Zhang, Bates, & Simon, 2011).

Medical homes rely heavily on the data derived from clinical decision support systems. These are information systems that provide the data needed for clinical staff to analyze patient diagnosis, disease management, illness prevention, and treatment. Decision support systems have been found to be effective in health care organizations that utilize them as a computerized tool. Gurwitz and Wolfstadt (2008) discovered that orders guided by decision support systems were associated with a 90 percent reduction in adverse drug events. The medical home benefits greatly from the system's ability to produce this outcome. In addition, the decision support system can play the role of sending computerized alerts to both the provider and the patient, such as providing notification that a mammogram is due (Gurwitz & Wolfstadt, 2008).

Electronic registries offer clinical staff at medical homes the ability to manage populations with chronic conditions. One technological tool is the Registry Population Manager. This system enables nurses and other clinical staff to identify populations of patients with condition specific criteria. Once the population is identified, the PCMH staff can perform intervention techniques, such as mailings or callings (Fleurant, et al., 2011; Zai, Estey, Lester & Yee, 2008).

Transition of care is important to the PCMH as patients are discharged from acute or post-acute care facilities. Patients should receive post discharge follow up care and medication management from the health professional that typically would be in the form of the patient's personal physician in a medical home setting. Information systems that reconcile a patient's post-discharge and pre-admit electronic medication listing have been evaluated for effectiveness. The tool should allow for smooth transition for a patient's care from the admit facility to the medical home (Bates & Bitton, 2010).

Telehealth is the use of modern telecommunication and electronic information to enable long distance care from the clinician to the patient. Such technologies include video conferencing, email communication, and electronic image transfer. Texas curbs spending by $1 billion by deploying EHRs and telehealth in prisons for the past decade. A telemedicine systems allowed Texas prisons to access healthcare providers at The University of Texas Medical Branch at Galveston in Galveston and Texas

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Tech University in Lubbock (iHealthBeat, 2011). Telehealth was valued by The Center for Information Technology Leadership at a projected savings across the U.S. of $4.3 billion annually (Bates & Bitton, 2010). Thus, medical homes are at a financially incentivized advantage to adopt telehealth as an option for care. The PPACA's support for the Patient Centered Medical Home is an example of innovation in modern health care. The PCMH's quality of care delivery system should be guided by technological advancements (Krupinski, 2008).

THE ELECTRONIC HEALTH RECORD CONCEPT

In 2009, under the Health Information Technology for Economic and Clinical Health (HITECH) act, the U.S. Department of Health & Human Services released proposed regulations that would support the use to the Electronic Health Record (EHR) in the health care industry. Health care providers would benefit from additional Medicare and Medicaid reimbursement with approved use of the EHR in their practices and facilities. Additional funding for users of EHRs would be provided as well (Blumenthal, 2010).

Ultimately the HITECH act's purpose was to improve care for citizens of the United States and technology was seen as a key component. Laying the groundwork for the EHR adoption would become a guide for health care systems, stand-alone facilities and other providers to use technology in order to optimize care effectively. Providers who could be recognized as "meaningful users" would be incentivized to implement systems that would qualify. The meaningful users of qualifying EHR systems would be required to show use of systems with the following characteristics: electronically capture coded health information, track key clinical conditions, share the information with other providers and patients, and report service quality measures of clinical data and public health information. Proof of these characteristics would help ensure the government's sound investment in the concept (Najaftorkaman, et al., 2015).

The EHR feeds life into a new approach to health care where technology is embraced. Records are kept electronically in virtually all industries and health care, though late in the game, has followed accordingly with the EHR.

The Paper-Based Medical Record

Hospitals and other health care providers in United States have historically kept paper-based medical records of patient files. The information is predominantly hand written which can cause issues of misinterpretation. Generally, paper files can only be stored at the physical location where the patient received the care. In fact, paper files of a patient's medical history are scattered across several locations – from the office of the primary care physician to the specialist's office to the central medical filing stations at the patient's acute care facility. If a physician requires medical history to treat a patient, other providers must be contacted to share this information. The transfer of this information is both time consuming and costly as files are passed on via fax, scan, or mail. In an industry where historical information is essential to provide the best possible care, modern technology possesses the cure for this dated system of paper filing (Schoen, OsbornSquires, Doty, Rasmussen, Pierson, & Applebaum, 2012).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The Electronic Medical Record (EMR)

An alternative to the paper file of a Care Delivery Organization (CDO) for any one particular patient is the Electronic Medical Record (EMR). The EMR is an electronic account of the patient encounter at the CDO. It contains clinical data, decision support, pharmacy data, and more across inpatient and outpatient environments of the CDO. It is important to note that the EMR does not store data from a patient's encounters at other CDO's, yet it serves as the basis for integrating a centralized record which will be discussed (Holroyd-Leduc, Lorenzetti, Straus, Sykes, & Quan, 2011).

The EMR's environment is the Clinical Data Repository (CDR). The CDR is referred to as a live database of patient information available to all practitioners within a CDO. Controlled Medical Vocabulary (CMV) is stored in the database and is pertinent to the patient encounter. This information must be accurately represented in the database in order for the Clinical Decision Support (CDS) and workflow of the EMR to perform as expected (Hillestad, Bigelow, Bower, Girosi, Meili, Scoville, & Taylor, 2005). The EMR components include clinical documentation, Computerized Provider Order Entry (CPOE), and pharmacy management (Garets & Davis, 2006).

The Electronic Health Record (EHR)

The Electronic Health Record (EHR) is reliant on the EMR being in place first. The EHR represents the environment where multiple stakeholders have access to a patient's medical information. The stakeholders include patients, providers, employers, payers (insurers), and the government. The stakeholders find value in obtaining this information to gain an advantage in accomplishing their purpose. While the EMR is owned by one singular CDO, the EHR is electronically shared and contains data of multiple CDO's within a community, region, state or even an entire country. National Health Information Network (NHIN) standards establish exchanges of data between the EMR to the EHR environment. The EHR environment consists of the stakeholder's and CDO's that rely on the data (Fontenot, 2014; Garets & Davis 2006).

EHR Advantages

Hospitals should predict beneficial outcomes that can be accomplished with the adoption of an EHR system. Advantages of EHR systems are exemplified by the hospital's ability to use the system in order to maintain preventive care measures, reduce operational expenses, pave the way for other forms of technology, and enhance access to data for hospital based research. EHRs can improve adherence rates for routine care and preventive screenings. Electronic alerts notify providers that a patient is due for a preventive screening. The alert system streamlines a process that would have normally been tedious and difficult to organize. The process is necessary and beneficial to the patients' health and well-being as preventive care is essential (Fontenot, 2014).

Treatments have become more effective with the use of EHRs. More patient information is available and easier to access by the nursing staff. One study found a 5% decrease in pressure ulcers after six months of an EHR system implementation (Menachemi & Collum, 2011). Menachemi and Collum (2011) found that the alert system, which is a part of Clinical Decision Support, accounted for an 11.3% appropriate hypertension treatment in primary care.

An initiative in the health care industry in recent years is to reduce "waste." Waste comes in the form of medical errors causing inappropriate costs, using less expensive alternatives of care, or fraud. Un-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

necessary laboratory and radiology tests are reduced with the presence of an EHR. This allows for the reduction in ancillary expense and the opportunity to fill the time saved performing inappropriate care with appropriate care (Menachemi & Collum, 2012).

An EHR capability that reduces errors in hospital operations is the Computerized Physician Order Entry (CPOE). A CPOE is an information system that is used in the medical field to communicate orders for a patient by a medical practitioner. The practitioner can electronically enter medication instructions, laboratory tests, radiology examinations and physical therapy orders. Prior to the development of this system, orders were written on paper and delivered to nursing, pharmacy and other medical staff. The existence of the CPOE system increases efficiency and reduces errors (Kaushal, Shojania, & Bates, 2003).

Operating expenses are reduced with the placement of an EHR system. The cost of chart pulls is virtually eliminated due to direct accessibility by the end user. Paper costs, previously used for creating paper based files, become almost obsolete. Salaries and wages can be reduced due the reduction in staff needed to perform duties that can be performed by the EHR. Legal expenses reduce as a result of increased communication across providers, adherence to medical guidelines, and reduction in malpractice claims (Mekhjian, Kumar, Kuehn, Bentley, Teater, Thomas, Payne, & Ahmad, 2002).

Hospitals that house research departments benefit from the EHR in that electronically stored data makes it more feasible for researches to collect information needed for studies. Measuring outcomes across a population in particular studies can be performed with ease because EHRs store data across multiple providers amongst numerous regions. The EHR advantages justify the effectiveness of the system. If the systems are implemented and maintained adequately, positive outcomes should be obtained.

EHR Limitations

There are several limitations of the EHR that have been recognized in the health care industry. These limitations include financial issues, changes in workflow, loss in productivity, threats to Protected Health Information (PHI) and others. The costs of implementing an EHR system include software and hardware purchases, overhead and additional salaries associated with converting existing paper based files to electronic, and training expenses for hospital staff. These costs affect both inpatient and outpatient settings in a health care organization. Ongoing maintenance costs include expenses related to regular updates and hardware replacement, salaries and benefits paid to hired technical staff to support the system, and fees paid to consultants and external contractors. Positive financial outcomes of an effective EHR system are shared by the third party insurance payer. The shared savings impact the payer as errors are eliminated and efficiency is improved from a billing standpoint. With the exception of Medicare, the payer does not invest financial resources in the adoption of the EHR. The entire cost is incurred by the health care provider. Thus, the benefit is reaped by the payer.

Loss in productivity can come in the form of medical staff spending non-productive time learning the system as opposed to providing medical care. EHR systems require extensive training for medical staff. Nurses, laboratory and radiology technicians, and other hospital staff are removed from their duties to learn the system. This process can take three months where a portion of the each staff member's worked hours are required for training each month. The productivity can be back filled with replacement staff. However, the replacement factor also leads to additional compensation, sometimes at a much higher hourly rate, which can burden the hospital budget. Because, hospital administrators emphasize the importance of keeping the Full Time Equivalent (FTE) budget on target, replacement can account for an unfavorable option.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

EHR's do not offer full protection in terms of patient privacy. As sensitive data is shared electronically over a network, Protected Health Information (PHI) data is at risk of IT security breaches just as credit card data is at risk in the retail industry. Regulations are required for hospital staff with access to PHI. Employees must adhere to the strong guidelines of the Health Insurance Portability and Accountability Act (HIPAA). This represents another challenge for health care organizations that must educate their staff on HIPAA regulations and how they are connected with EHR systems (U.S. Department of Health & Human Services, 2015).

Other limitations and setbacks of the technology can include an overdependence on the technology, increases medical errors if the system is poorly designed or if the employees are not adequately trained, and loss of autonomy for physicians due to system limitations. The limitations can be addressed with strategic planning by hospital administrators. Also, as systems become more mature, limitations are liable to be alleviated.

An EHR Study: Costs and Benefits Analysis

Gundersen Lutheran Medical Center in La Crosse, Wisconsin, was examined in a study of the outcomes demonstrated by one hospital that adopted the EHR. The hospital is a 325 licensed bed trauma center that implemented an inpatient EHR system in November of 2008. For comparison purposes, data were collected one year pre-EHR implementation and one year post-EHR implementation. For purposes of this study, the impact of the CPOE system implementation was considered in the results because of the proximity of the CPOE implementation to that of the EHR and for the reason that CPOE and EHR systems are directly linked in terms of their capabilities (Zlabeck, Wickus, & Mathiason, 2011).

Laboratory tests and radiology examinations were analyzed in the study. The analysis showed that the number of laboratory test per week per hospitalization reduced from 13.9 in the pre-EHR period data to 11.9 in the post-EHR data. Zlabeck, Wickus, and Mathiason (2011) point out that this decrease of 18% in the rate of lab tests begin to show signs of significant trend at month three of the post-EHR period. This leads to the suggestion that the reduction in tests was a result of the CPOE system's implementation. However, the EHR needed to be in place in order for the CPOE system's implementation to occur. Therefore, both systems should be recognized in this performance driven statistic. Similarly, the rate of radiology tests per hospitalization did not show signs of a decrease until month three of the post-EHR period. The number decreased from 2.06 in the pre-EHR period to 1.93 in month 9 of the post-EHR period.

Significant savings were demonstrated in transcription costs. The cost per month decreased from $74,596 to $18,938. This represented an impressive outcome where costs reduced by 74.6% from the pre-EHR period to the post-EHR period. Over the twelve month post-EHR period, total savings calculated to $667,896. This outcome was attributable to the EHR system directly (Zlabeck, Wickus, and Mathiason, 2011).

An obvious positive financial outcome came in the form of cost savings in supplies specifically related to a reduction in paper consumption. The EHR eliminates the need for paper based medical files, therefore this cost savings would generally be expected. The hospital experienced a 26.6% decrease in reams of paper orders, and one hundred and eighteen paper forms were eliminated. The total cost savings decreased in the post-EHR period by a total of $30,351. Medication errors were reduced after the EHR implementation as well. Errors per 1,000 hospital days dropped from 17.9 prior to the EHR system to 15.4 by month nine in the post-EHR period. As was the case with lab tests and radiology exams, the CPOE played a vital role in this outcome (Mekhjian, et al., 2002; Zlabeck, Wickus, & Mathiason, 2011) .

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The statistics prove that positive operational and cost benefits were associated with the implementation of the EHR and CPOE systems at Gundersen Lutheran Medical Center. The study also demonstrates that the adoption of the EHR can open doors to other modern formers of Health Information Technology. In this case it led to an effective implementation of the CPOE system (Mekhjian, et al., 2002).

Epic: Judith Faulkner's Systems

Judith Faulkner has built Epic Systems gradually and quietly over several years of development, capitalizing off of an evolving health care environment that has recently adapted to the benefits the EHR. Epic was founded in 1979 in Verona, Wisconsin and is presently becoming the preferred technological system solution for the majority of the stakeholders in the U.S. health care industry. Its clients include dominant health care systems such as Johns Hopkins and Cleveland Clinic (Moukheiber, 2012). Massachusetts-based Partners HealthCare, founded by the prestigious Massachusetts General Hospital and Brigham and Women's Hospital, is currently implementing the system across all of its hospitals – the first of which began in February of 2014 at Newton Wellesley Hospital. In an article in Forbes Magazine, it was determined that 40% of the U.S. population is being stored in an Epic digital record. This share continues to grow as the U.S. health care industry proceeds to become more innovative in the field of technology (Moukheiber, 2012).

Epic develops software for medical groups, hospitals and integrated health care systems. At a lower level, it serves community based hospitals, academic facilities, children's hospitals and health safety net providers (Ku, et al., 2009). The EHR methodology employed by Epic is based on a health care programming language referred to as MUMPS (Massachusetts General Hospital Utility Multi-Programming System). Its products include EpicCare EMR that normalizes care delivery for physicians, EpicCare Inpatient Clinical System that allows for health information sharing across organizations, mobile applications, and portals. Also available is a resource for patients to retrieve the same electronic medical records that are accessed by their physicians. A patient's Personal Health Record (PHR) is portable for both patients and providers alike through an Epic application known as Lucy. Epic also offers PlanLink which is a tool for payers (i.e., insurers) and employers that makes information available via the web (Bloomberg Business Week, 2014; Holroyd-Leduc, et al., 2011; O'Brien, 2006;).

In the era of digitalized health care, security is a strong selling point when it comes to investing in health care IT. Faulkner backs up her product's security status by proclaiming "to the best of our knowledge, in the 32 years we've been in business, there has never been a breach of Epic's data by a hacker" (Freudenheim, 2012). The company's leader stresses the importance of firewalls and security systems that Epic has in place to protect sensitive and protected data. In fact, Faulkner holds the position of an industry representative on a government panel that examines security issues in health care. Her goal is to assure the protection of the data while still allowing for the flow of information sharing. Such a balance is difficult to reach with success, yet Epic manages to do so strategically. The information sharing can only be shared across providers if they are clients or if they meet certain conditions. Organizations must prove patient consent and have a digital certificate to be presented to other organizations in order to engage in digital information sharing. The benefit of sharing health care data is something the Faulkner sees as a privilege that should be treated accordingly. The Health Insurance Portability and Accountability Act of 1996 mandates the privacy and security of sensitive patient data. Epic's reputation for safeguarding the data increases the effectiveness and marketability of its software products (Freudenheim, 2012; O'Brien, 2006).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

USE OF HEALTH CARE IT TO IMPROVE SERVICE QUALITY

Healthcare IT applications in the medical community bring in many strengths of service quality. The strengths, broadly speaking, include user satisfaction, employee productivity, and business performance. Weaknesses include stakeholder issues, regulatory environment, technical issues, and financial issues. Threats include government mandates, third party e-Health and e-Medicine, competition for scarce IT resources and talent, market and competitive pressure, and the threat of government run healthcare. There are opportunities as well. Institutions that are developing effective systems now have the benefit of being early adopters. They can affect change of standards toward those that are beneficial to them since these standards are just now being developed. They can appeal to younger patients searching for locations and healthcare professionals with better service quality. They can gain prestige among peers and increase their own service quality outcomes (Andrews, 2015; Lettieri, Bartoli, & Masella, 2013).

The strength of user satisfaction takes a couple of different forms. First there are now more Internet interactions than ever before. People are used to doing their banking, shopping for products, arranging food deliveries, paying bills, communicating, taking classes, and many other aspects of daily living online. They are often surprised that medical processes are still inaccessible in many facets. This has led to higher patient expectation that can be met with Health IT solutions. People are more tech savvy now as well that makes a more educated and able consumer. People are also beginning to shop for healthcare services and products including both physicians and facilities (Matysiewicz & Smyczek, 2013; Razzaque, Eldabi, & Jalal-Karim, 2013).

The strength of employee productivity is due to labor efficiency and service effectiveness. There is a shortage of health care professionals especially in the general practitioner area. Manual record keeping and searches are a laborious and time consuming process. Health IT can deliver instant information on demand to improve labor efficiency. The accuracy of health IT generated reports will improve the diagnosis and service effectiveness. Employee productivity can be increases by using health IT applications to free up data searching time and to help serve the growing patient need. The strength of business performance centers on using health IT to reduce costs. Health IT applications of efficiencies increase employee productivity and reduce operating costs. Health medical records are stored electronically and can be queried or reported in various customized ways. Business performance can be further increased by reducing malpractice through automatic guards, checks, and alerts (Menachemi, & Collum, 2011; Menon & Kohli, 2013).

Information Systems Guide Management in Health Benefits and Costs

In 2013, employer sponsored health insurance covered about 149 million employees and their dependents. Over the past decade, employer sponsored health insurance costs have trended at a rate much higher than inflation. Data from the 2003 and 2013 versions of the Kaiser Survey of Employer Sponsored Health Benefits depict an 80% increase in the average total insurance premiums and an 89% increase in the average employee contribution to the full premium. In 2003, employers subsidized the full health insurance premium by a very generous 75%. However, as the employee contributions have experienced an upward trend at a higher rate than the total, the employer sponsored subsidy has decreased over the ten year period to 74% (Kaiser Family Foundation, 2013).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The data infer that in an effort to withstand the astronomical spike in health benefit costs, employers have decided to cut benefits and push a higher share of the premium payments onto their employees. As companies seek to both cut costs and remain competitive from a compensation standpoint, they have invested in data warehouses such as Truven Health Analytics and other systems. The investment in these systems is a practice aimed towards organizations empowering themselves with the reporting capabilities needed to analyze data and develop strategies for mitigating the trend of health insurance costs (Vo & Bhaskar, 2012).

Truven Health Analytics offers a product to employers called Advantage Suite, which is a dynamic database that stores in-depth health and Rx claim data of employees and their dependents. Member information is submitted to Truven by the group health and Rx insurance carriers that the employer uses to offer health benefits to its employees. The employer submits job and demographic data that are combined to the claims data by Truven to build a database the employers can use to report on clinical and financial member data. Advantage Suite allows employers to dive into the data of their health benefit costs and utilization. Employers can determine the drivers for substantial increases in trend and develop plans of action to reduce costs. They can identify preventive measures, such as colonoscopy and mammography rates, in their population and compare these measures to norms using a built in benchmarking tool. The tool pulls information using a resource known as MarketScan data, a comparative application that can be industry and demographically adjusted. After interpreting the results, employers can take action to steer their population into healthy behavior and ultimately reduce costs or care related to chronic conditions. Employers can also use the system to develop forecast and budget assumptions on employee health expenses and perform benefit modeling analytics used to measure the cost impact of plan changes.

In a study performed by the Employee Benefit Research Institute (EBRI), Truven data were used to examine the costs of spouses in employer sponsored health plans. The EBRI sought to address the notion that spouses incur higher health costs in employer sponsored health plans. Using Truven Health Analytics MarketScan Commercial Database, researchers were able to analyze claims data of 317,180 couples and 634,360 individual members across multiple employers. The system allowed the researchers to dice the data by geographic location, health status, gender, and age. For the population analyzed, the results showed that the employees (i.e., policy holders) cost $5,430 on average while the spouses averaged $6,609. The findings revealed a 21.7% increase of the spouse cost per member over that of the subscribing employee. A powerful information system allows the end user to uncover the layers by drilling down into the detail to determine the root causes.

By running subsequent reports, Fronstin and Roebuck (2014) found that the majority of the $1,179 difference was due to gender, health status, and other factors. Researchers adjusted their reports to normalize the data and still found that average spouse costs were 7% higher than the average policy holder costs. The end result of the analysis was that, even when adjusting for outliers, spouses still carry a higher cost on average than the policy holding employee.

If an employer performs a similar analysis using a Truven product or a similar system, decisions can be made to reduce employee health costs. The EBRI indicates that the United Postal Service (UPS) applied the same approach to reduce health costs. UPS decided to change its benefit package by only offering coverage to the spouses of its employees who did not have coverage offered to them by their own employer. UPS actually cited the PPACA as the cause of its decision, indicating that the change enabled the company to provide affordable coverage. The PPACA requires employers to offer coverage to dependent children through the age of 26, but does not require coverage to be offered to spouses (Fronstin & Roebuck, 2014).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Were it not for a sophisticated technologically advanced system like Truven's, the analysis would have limitations and perhaps UPS would not be able to place a value on the cost impact of its decision. The investment in information technology by employers can prove to carry a positive return, because the accessible data empowers employers with knowledge that would otherwise be unavailable.

Information Systems Guide Employee Wellness and Productivity

A healthy workforce translates into a productive workforce. In order to remain competitive, an organization must take action in order to keep its workforce healthy and productive. Information systems, like Truven's, allow employers to track clinical measures on employee health to identify areas that can be improved upon. Truven's Advantage Suite data warehouse offers a reporting environment that allows employers to track the effectiveness of disease management programs.

Insurers often enroll members in these disease management programs, which are programs equipped with clinical staff who offer health coaching to members with certain clinical conditions. Using the condition of diabetes as an example, a member with the disease would generally be contacted with reminders to get their periodic exams such as HbA1c and eye tests needed to maintain the well-being of the member. An employer with Truven Advantage Suite could access the reporting environment, set their clinical condition criteria to diabetes, and query on the frequency of periodic exams and the overall health of their members. Because the database offers time constraints, the employer can track the improvement over multiple time periods to analyze trends. If the employer wishes to set parameters on the population within the clinical condition population, subsets can be created limiting the person IDs. The use of subsets is most effective when the analysis calls for an examination on cohorts. If issues are identified, such as an unfavorable rate in HbA1c testing, employers can deliver the results to the insurer and order further intervention (Truven Health Analytics, 2014).

An employer can also use the database to perform home grown analytics on the overall health and wellness of the employee population. For instance, through statistical analysis a company may reveal a portion of their population to have a higher prevalence for mental health conditions, such as depression. That company might decide to establish an employee assistance program for these employees in order for them to have access to counseling on site. HIPAA complicates the communication process, however, if done strategically the wellness initiative can foster a healthy workforce.

IT PROFESSIONS IN HEALTH CARE

Role of the Chief Medical Information Officer (CMIO)

The Chief Medical Information Officer (CMIO) role has recently been gaining popularity in the health care industry. The CMIO is responsible for directing the implementation and maintenance of clinical information systems. Generally, this position is held by a physician. The emergence of the EHR and EMR has added depth to the CMIO role as health care organizations convert to digital record keeping. Responsibilities also include facilitating data warehouse initiatives, physician website and email maintenance, and network security. CMIOs can be partially dedicated to performing direct patient care as attending physicians or fully dedicated to their roles in health care IT (Holroyd-Leduc, et al., 2011).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CMIOs set goals and expectations for IT projects. They also lead the vendor selection process when new systems or technology are requested. For instance, the CMIO would be responsible for the selection process involved for a health care organization to choose Epic Systems to replace an existing system, such a Meditech. Once implemented, the CMIO would then direct plans for training registered nurses and other clinical staff on new information. CMIOs can also chair senior health system committees and information technology committees. Also, CMIOs participate and direct the establishment of information security and that HIPAA regulations are carried out from a systems standpoint. CMIOs lead staff reporting directly to them or collaborate with the organizations internal IT staff to achieve the effective completion of these initiatives (Leviss, Kremsdorf, & Mohaideen, 2006).

Nursing Informatics

The American Medical Informatics Association (AMIA) defines nursing informatics as the "science and practice (that) integrate nursing, its information and knowledge, with management of information and communication technologies to promote the health of people, families, and communities worldwide" (American Medical Informatics Association (AMIA), 2014). Positions in this field are staffed by health care organizations and include Nurse Informaticists and Chief Nursing Informatics Officers. Nursing informatics is expanding and the need for these positions is increasing. The 2014 Healthcare Information and Management Systems Society (HIMSS) Nursing Informatics Survey revealed that there are three areas where nursing informatics add value to an integrated health care system. These areas are the development and implementation of IT systems, optimization of HIT systems, defining context for patient care. The survey indicates that 30 percent of health care organizations support the CNIO position. The role is evolving as that number is expected to grow in subsequent years (Gabriel, Jones, Samy, & King, 2014; Pizzi, 2014).

Clinical Decision Support (CDS)

Clinical Decision Support (CDS) systems provide intelligence that fuels decision making in the health care environment. These decisions benefit clinical staff and patients, enhance the delivery of care, control costs, and increase revenue. CDS system analysts filter information to the health care environment using technical tools such as computerized alerts, reporting platforms, diagnostic support, and clinical guidelines. Ultimately, health care organizations improve from service quality decision support by increasing the quality of care, reducing errors and complications, improving efficiency, and satisfying patients (Kaushal, Shojania, & Bates, 2003; Romano & Stafford, 2011).

Health care organizations staff departments specifically dedicated to CDS. These departments tend to consist of analysts who operate and maintain databases used to provide support for the decisions made within their organizations. One of these systems is Allscripts EPSi. EPSi is a web-based information system that is specific to the health care industry and stores data and analytical tools that guide financial management. CDS analysts operate the system to provide integrated analytics, budget support, and general decision support. EPSi offers multiple modules to organizations that include Budget Manager, Capital Budget Manager, Strategic Product Budgeting, Productivity Manager, Cost Manager, Product Line Analyst, and EPSi Management Dashboard. Budget Manager allows the health care organization to combine payroll and capital budgeting with patient level budgeting so that realistic spending goals and assumptions can be made. Analyst can perform financial modeling using this module's ad hoc report writing resource and flex budget mode (Allscripts, 2014).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Analysts in health care organizations use IT resources, such as EPSi, in order to equip themselves and their organizations with the information needed to determine appropriate investment and allocation of resources. The information system makes this possible. Without the assistance of EPSi and other CDS systems, analysts would not be capable of impacting the health care system. CDS systems have paved the way for business intelligence professions in health care organizations.

CONCLUSION

Health care systems continue to swell into massive heaps of facilities, insurance companies, service organizations, consultants and more. Partners HealthCare consists of entities from acute care facilities to mental health facilities to rehabilitation hospitals and centers. Partners and other Massachusetts based systems, like Steward Health Care and Hallmark Health, base their business models on the concept of integrating care across their systems. IT has played in integral role in effectively coordinating care within any one particular system. What remains to be seen is how effectively IT will integrate all U.S. health care organizations. The Electronic Health Record will be the element that allows organizations to facilitate with one another from a technical sense. Patient information must be properly shared for the betterment of patient care and effectiveness of care delivery by health care professionals. As the adoption of the EHR spreads across the United States in future years, systems will improve their care delivery models and processes.

The government has established itself as a major player in the industry with the passing of the PPACA. President Obama's Health Care Reform has been designed to steer the industry improvement and adoption. As more focus is placed on IT resources within health care organizations, systems should continue to improve and IT departments should expand.

Epic will continue to sweep the nation with its state of the art systems catered to the current need for today's health care environment. Systems that previously controlled the market, such as Soarian and Meditech, will need to improve their products as soon as possible in order to withstand Epic's domination. The emergence of a new entrant into the world of IT systems will be a more likely scenario.

The access to systems will expand for organizations who offer employer sponsored health plans to their employees and the families of their employees. Systems, like Truven, will be adopted by both employers and benefit consultants across the U.S. The return on investment has been demonstrated by the early adopters of these IT systems and more employers will fall in line to cut costs and keep their workforce healthy and productive.

Health care organizations can be for profit or non-profit, acute or non-acute, outpatient or inpatient. Regardless of an organization's disposition, it cannot continue to exist without information sharing and gathering through technological advancements. IT analytical tools provide the business intelligence needed for health care organizations to survive and flourish. It is with these tools that health care professionals strive to fulfill their duties and work for the betterment of the patient in their care.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Allscripts. (2014). *Solutions that round-out all the elements of financial health*. Retrieved December 1, 2014 from https://www.allscripts.com/en/solutions/acute-solutions/performance-management/performancemanagement.html

American Medical Informatics Association (AMIA). (2014). *Nursing Informatics*. Retrieved December 1, 2014 from http://www.amia.org/programs/working-groups/nursing-informatics

Andrews, R. M. (2015). Statewide hospital discharge data: Collection, use, limitations, and improvements. *Health Services Research*, *50*(1), 1273–1299. doi:10.1111/1475-6773.12343 PMID:26150118

Bates, D. W., & Bitton, A. (2010). The future of health information technology in the patient-centered medical home. *Health Affairs*, *29*(4), 614–621. doi:10.1377/hlthaff.2010.0007 PMID:20368590

Berwick, D. (2011). Launching Accountable Care Organizations — the proposed rule for the Medicare shared savings program. *The New England Journal of Medicine*, *10*(1056), 1–4. PMID:21452999

Bitton, A., Filer, L. A., & Jha, A. K. (2012). Health information technology in the era of care delivery reform. *Journal of the American Medical Association*, *307*(24), 2593–2594. doi:10.1001/jama.2012.6663 PMID:22735426

Bloomberg Business Week. (2014). *Company overview of Epic Systems Corporation*. Retrieved December 10, 2014 from http://investing.businessweek.com/research/stocks/private/snapshot.asp?privcapId=6479863

Blumenthal, D. (2010). Launching HITECH. *The New England Journal of Medicine*, *362*(5), 382–385. doi:10.1056/NEJMp0912825 PMID:20042745

Carey, M. A., & Appleby, J. (2013). *How do the health law's online exchanges work? What you need to know to get started.* Retrieved December 10, 2014 from http://www.washingtonpost.com/national/health-science/how-do-the-health-laws-online-exchanges-work-what-you-need-to-know-to-get-started/2013/09/16/ef7974c8-1988-11e3-a628-7e6dde8f889d_story.html

Epstein, R. M., Fiscella, K., Lesser, C. S., & Stange, K. C. (2010). Why the nation needs a policy push on patient-centered health care. *Health Affairs*, *29*(8), 1489–1495. doi:10.1377/hlthaff.2009.0888 PMID:20679652

Fleurant, M., Kell, R., Love, J., Jenter, C., Volk, L. A., Zhang, F., ... Simon, S. R. (2011). Massachusetts e-Health Project increased physicians' ability to use registries, and signals progress toward better care. *Health Affairs (Project Hope)*, *30*(7), 1256–1264. doi:10.1377/hlthaff.2010.1020 PMID:21734198

Fontenot, S. F. (2014). The affordable care act and electronic health care records: Can technology help reduce the cost of health care? *Physician Executive*, *40*(1), 68–72. PMID:24575706

Ford, P. (2013). *The Obamacare website didn't have to fail. How to do better next time*. Retrieved December 10, 2014 from http://www.businessweek.com/articles/2013-10-16/open-source-everything-the-moral-of-the-healthcare-dot-gov-debacle#p3

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Freudenheim, M. (2012). Digitizing health records before it was cool. *The New York Times*. Retrieved December 10, 2014 from http://www.nytimes.com/2012/01/15/business/epic-systems-digitizing-health-records-before-it-was-cool.html?pagewanted=all&_r=1&

Fronstin, P., & Roebuck, C. (2014). The cost of spousal health coverage. *Employee Benefit Research Institute*, *35*(1), 2-7. Retrieved December 1, 2014 from http://www.ebri.org/pdf/notespdf/EBRI_Notes_01_Jan-14_SpslCvg-RetPlns1.pdf

Gabriel, M. H., Jones, E. B., Samy, L., & King, J. (2014). Progress and challenges: Implementation and use of health information technology among critical-access hospitals. *Health Affairs*, *33*(7), 1262–1270. doi:10.1377/hlthaff.2014.0279 PMID:25006155

Garets, D., & Davis, M. (2006). Electronic medical records vs. electronic health records: Yes, there is a difference. *HIMSS Analytics*. Retrieved December 10, 2014 from http://www.himss.org/files/HIMSSorg/content/files/WP_EMR_EHR.pdf

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. E. M. R. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs*, *24*(5), 1103–1117. doi:10.1377/hlthaff.24.5.1103 PMID:16162551

Holroyd-Leduc, J., Lorenzetti, D., Straus, S. E., Sykes, L., & Quan, H. (2011). The impact of the electronic medical record on structure, process, and outcomes within primary care: A systematic review of the evidence. *Journal of the American Medical Informatics Association: JAMIA*, *18*(6), 732–737. doi:10.1136/amiajnl-2010-000019 PMID:21659445

iHealthBeat. (2011). *Texas curbs spending by $1B by deploying EHRs, telehealth in prisons*. Retrieved December 1, 2014 from http://www.ihealthbeat.org/articles/2011/8/26/texas-curbs-spending-by-1b-by-deploying-ehrs-telehealth-in-prisons

Kaiser Family Foundation. (2013). Employer Health Benefits: 2013 Summary of Findings. *The Kaiser Family Foundation and Health Research and Educational Trust*. Retrieved December 1, 2014 from http://kaiserfamilyfoundation.files.wordpress.com/2013/08/8466-employer-health-benefits-2013_summary-of-findings2.pdf

Kaushal, R., Shojania, K. G., & Bates, D. W. (2003). Effects of computerized physician order entry and clinical decision support systems on medication safety: A systematic review. *Archives of Internal Medicine*, *163*(12), 1409–1416. doi:10.1001/archinte.163.12.1409 PMID:12824090

Krupinski, E. A. (2008). Telemedicine for home health and the new patient: When do we really need to go to the hospital? *Studies in Health Technology and Informatics*, *29*(4), 614–621. PMID:18305331

Ku, L., Jones, E., Finnegan, B., Shin, P., & Rosenbaum, S. (2009). *How is the primary care safety net faring in Massachusetts? Community health centers in the midst of health reform (Issue paper no. 7878)*. Washington, DC: Kaiser Commission on Medicaid and the Uninsured.

Lettieri, E., Bartoli, L., & Masella, C. (2013). Coordinating intra-sector services in healthcare: Requirements and elements that managers should take into account. *European Management Journal*, *31*(6), 591–601. doi:10.1016/j.emj.2013.05.004

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Leviss, J., Kremsdorf, R., & Mohaideen, M. (2006). The CMIO – A new leader for health systems. *Journal of the American Medical Informatics Association*, *13*(5), 573–578. doi:10.1197/jamia.M2097 PMID:16799119

Matysiewicz, J., & Smyczek, S. (2013). Relations between healthcare organizations and their patients: Three-factors models in the e-healthcare market. *Review of Business*, *33*(2), 7–25.

Mekhjian, H. S., Kumar, R. R., Kuehn, L., Bentley, T. D., Teater, P., Thomas, A., ... Ahmad, A. (2002). Immediate benefits realized following implementation of physician order entry at an academic medical center. *Journal of the American Medical Informatics Association*, *9*(5), 529–539. doi:10.1197/jamia.M1038 PMID:12223505

Menachemi, N., & Collum, T. (2011). Benefits and drawbacks of electronic health record systems. *Risk Management and Healthcare Policy*, *4*, 47–55. doi:10.2147/RMHP.S12985 PMID:22312227

Menon, N. M., & Kohli, R. (2013). Blunting Damocles' Sword: A longitudinal model of healthcare IT impact on malpractice insurance premium and quality of patient care. *Information Systems Research*, *24*(4), 918–932, 1164–1165. doi:10.1287/isre.2013.0484

Moukheiber, Z. (2012). *Epic systems' tough billionaire*. Retrieved December 10, 2014 from http://www.forbes.com/sites/zinamoukheiber/2012/04/18/epic-systems-tough-billionaire/

Najaftorkaman, M., Ghapanchi, A. H., Talaei-Khoei, A., & Ray, P. (2015). A taxonomy of antecedents to user adoption of health information systems: A synthesis of thirty years of research. *Journal of the Association for Information Science and Technology*, *66*(3), 576–598. doi:10.1002/asi.23181

Nussbaum, A., Tirrell, M., Wechsler, P., & Randall, T. (2010, April 5). Obamacare's cost scalpel. *Business Week*, 64-65.

O'Brien, M. S. (2006). Implementation of the EPIC Electronic Medical Record/Physician Order-Entry System. *Journal of Healthcare Management*, *51*(5), 338–343. PMID:17039692

Pizzi, R. (2014). Nursing Informatics Adds Value as Healthcare is Transforming. *Healthcare IT News*. Retrieved December 1, 2014 from http://www.healthcareitnews.com/news/nursing-informatics-adds-value-healthcare-transforming

Razzaque, A., Eldabi, T., & Jalal-Karim, A. (2013). Physician virtual community and medical decision making: Mediating role of knowledge sharing. *Journal of Enterprise Information Management*, *26*(5), 500–515. doi:10.1108/JEIM-07-2013-0047

Romano, M. J., & Stafford, R. S. (2011). Electronic health records and clinical decision support systems: Impact on national ambulatory care quality. *National Library of Medicine. Archives of Internal Medicine*, *171*(10), 897–903. doi:10.1001/archinternmed.2010.527 PMID:21263077

Rosenbaum, S., & Margulies, R. (2011). Tax-exempt hospitals and the Patient Protection and Affordable Care Act: Implications for public health policy and practice. *National Library of Medicine. Public Health Reports*, *126*(2), 283–286. PMID:21387959

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Schoen, C., Osborn, R., Squires, D., Doty, M., Rasmussen, P., Pierson, R., & Applebaum, S. (2012). A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. *Health Affairs*, *31*(12), 2805–2816. doi:10.1377/hlthaff.2012.0884 PMID:23154997

Somashekhar, S., & Sun, L. H. (2013). *HealthCare.gov meets deadline for fixes Obama administration says*. Retrieved December 10, 2014 from http://www.washingtonpost.com/national/health-science/healthcaregov-meets-deadline-for-fixes-white-house-says/2013/12/01/a3885612-5a13-11e3-ba82-16ed03681809_story.html

Thibodeau, P., & Gross, G. (2013). HealthCare.gov had "no chance in hell". *Computerworld*, *47*(19), 6.

Truven Health Analytics. (2014). *Disease Management Analytics*. Retrieved December 1, 2014 from http://truvenhealth.com/your-healthcare-focus/life-sciences/disease-management-analytics

U.S. Department of Health & Human Services. (2014). What is a medical home? Why is it important? *Health Resources and Services Administration (HRSA) - Health Information Technology*. Retrieved December 10, 2014 from http://www.hrsa.gov/healthit/toolbox/Childrenstoolbox/BuildingMedicalHome/whyimportant.html

U.S. Department of Health & Human Services. (2015). *Summary of the HIPAA security rule*. Retrieved April 20, 2015 from http://www.hhs.gov/ocr/privacy/hipaa/understanding/srsummary.html

Vo, A., & Bhaskar, R. (2012). Health Care Reform Requires Rethinking on the IT Strategy. *Journal of Cases on Information Technology*, *14*(2), 65–72. doi:10.4018/jcit.2012040105

Wolfstadt, J. I., Gurwitz, J. H., Field, T. S., Lee, M., Kalkar, S., Wu, W., & Rochon, P. A. (2008). The effect of computerized physician order entry with clinical decision support on the rates of adverse drug events: A systematic review. *Journal of General Internal Medicine*, *23*(4), 451–458. doi:10.100711606-008-0504-5 PMID:18373144

Zai, A. H., Grant, R. W., Estey, G., Lester, W. T., Andrews, C. T., Yee, R., ... Chueh, H. (2008). Lessons from implementing a combined workflow-informatics system for diabetes management. *Journal of the American Medical Informatics Association*, *15*(4), 524–533. doi:10.1197/jamia.M2598 PMID:18436907

Zlabek, J. A., Wickus, J. W., & Mathiason, M. A. (2011). Early cost and safety benefits of an inpatient electronic health record. *Journal of the American Medical Informatics Association*, *18*(2), 169–172. doi:10.1136/jamia.2010.007229 PMID:21292703

ADDITIONAL READING

Agarwal, R., Sands, D. Z., Schneider, J. D., & Smaltz, D. H. (2010). Quantifying the economic impact of communication inefficiencies in U.S. hospitals. *Journal of Healthcare Management*, *55*(4), 265–282. PMID:20812527

Alpay, L., Van Der Boog, P., & Dumaij, A. (2011). An empowerment-based approach to developing innovative e-health tools for self-management. *Health Informatics Journal*, *17*(4), 247–255. doi:10.1177/1460458211420089 PMID:22193825

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ball, M. J., & Lillis, J. (2001). E-health: Transforming the physician/patient relationship. *International Journal of Medical Informatics*, *61*(1), 1–10. doi:10.1016/S1386-5056(00)00130-1 PMID:11248599

Ballantine, P. W., & Stephenson, R. J. (2011). Help me, I'm fat! Social support in online weight loss networks. *Journal of Consumer Behaviour*, *10*(5), 332–337. doi:10.1002/cb.374

Bitner, M. J. (2001). Service and technology: Opportunities and paradoxes. *Managing Service Quality*, *11*(6), 375–379. doi:10.1108/09604520110410584

Bringewatt, R. J. (2001). Making a business case for high-quality chronic illness care. *Health Affairs*, *20*(6), 59–60. doi:10.1377/hlthaff.20.6.59 PMID:11816690

Bugshan, H., Nick Hajli, M., Lin, X., Featherman, M., & Cohen, I. (2014). Social media for developing health services. *Qualitative Market Research*, *17*(3), 283–296. doi:10.1108/QMR-07-2013-0048

Chou, W. Y. S., Prestin, A., Lyons, C., & Wen, K. Y. (2013). Web 2.0 for health promotion: Reviewing the current evidence. *American Journal of Public Health*, *103*(1), 9–18. doi:10.2105/AJPH.2012.301071 PMID:23153164

Cunningham, P., Grant-Pearce, C., Green, L., Miles, I., Rigby, J., & Uyarra, E. (2005). In Sickness, in Health and in Innovation: NHS DIRECT – A Health Sector Innovation Study. Administration, 53(3), 42–65.

Ghannouchi, S. A., Mabrouk, K., & Ghannouchi, S. (2010). Proposal of data warehouse in the context of healthcare process reengineering. *Business Process Management Journal*, *16*(4), 688–712. doi:10.1108/14637151011065964

Gibbons, M. C., Fleisher, L., Slamon, R. E., Bass, S., Kandadai, V., & Beck, J. R. (2011). Exploring the potential of Web 2.0 to address health disparities. *Journal of Health Communication*, *16*(1), 77–89. doi:10.1080/10810730.2011.596916 PMID:21843097

Groselj, D. (2014). A webometric analysis of online health information: Sponsorship, platform type and link structures. *Online Information Review*, *38*(2), 209–231. doi:10.1108/OIR-01-2013-0011

Hajli, M., Bugshan, H., Lin, X., & Featherman, M. (2013). From e-learning to social learning-a health care study. *European Journal of Training and Development*, *37*(9), 851–863. doi:10.1108/EJTD-10-2012-0062

Halamka, J. D. (2014). Early experiences with big data at an academic medical center. *Health Affairs*, *33*(7), 1132–1138. doi:10.1377/hlthaff.2014.0031 PMID:25006138

Hamid, A., & Sarmad, A. (2008). Evaluation of e-health services: user's perspective criteria. *Transforming Government: People. Process and Policy*, *2*(4), 243–255.

Hardey, M. (2001). 'E-health': The internet and the transformation of patients into consumers and producers of health knowledge. *Information Communication and Society*, *4*(3), 388–405. doi:10.1080/713768551

Harjumaa, M., & Muuraiskangas, S. (2014). Building persuasiveness into information systems. *Electronic Journal of Information Systems Evaluation*, *17*(1), 23–35.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kamel, B., Maged, N., & Wheeler, S. (2007). The emerging Web 2.0 social software: An enabling suite of sociable technologies in health and health care education1. *Health Information and Libraries Journal*, *24*(1), 2–23. doi:10.1111/j.1471-1842.2007.00701.x PMID:17331140

Liu, L., & Zhu, D. (2013). An integrated e-service model for electronic medical records. *Information Systems and eBusiness Management, 11*(1), 161-183.

Maletz, F. W. (2011). From hospital to "Healthspital": A better paradigm for health care. *The Futurist*, *45*(2), 16–19.

Neuhauser, L., & Kreps, G. L. (2003). Rethinking communication in the e-health era. *Journal of Health Psychology*, *8*(1), 7–23. doi:10.1177/1359105303008001426 PMID:22113897

Palvia, P., & Thambusamy, R. (2013). Electronic medical records application development: Perspectives of the service provider. *Journal of Information Technology Case and Application Research*, *15*(2), 11–29. doi:10.1080/15228053.2013.11082801

Phillips, A., & Thornbory, G. (2014). Apps for health professionals. *Occupational Health*, *66*(4), 27–29.

Schartinger, D., Miles, I., Saritas, O., Amanatidou, E., Giesecke, S., Heller-Schuh, B., ... Schreier, G. (2015). Personal health systems technologies: Critical issues in service innovation and diffusion. *Technology Innovation Management Review*, *5*(2), 46–57.

Sunyaev, A. (2013). Consumer facing health care systems. *E-Service Journal, 9*(2), 1-23, 63.

Wilson, E. V., & Lankton, N. K. (2004). Modeling patients' acceptance of provider-delivered e-health. *Journal of the American Medical Informatics Association*, *11*(4), 241–248. doi:10.1197/jamia.M1475 PMID:15064290

Wilson, E. V., & Tulu, B. (2010). The rise of a health-IT academic focus. *Communications of the ACM*, *53*(5), 147–150. doi:10.1145/1735223.1735259

This research was previously published in Effective Methods for Modern Healthcare Service Quality and Evaluation edited by Panagiotis Manolitzas, Evangelos Grigoroudis, Nikolaos Matsatsinis, and Denis Yannacopoulos; pages 110-136; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 98
Early Detection of Dementia: Advances, Challenges, and Future Prospects

Stefanos Xefteris
Aristotle University of Thessaloniki, Greece

Charis Styliadis
Aristotle University of Thessaloniki, Greece

Evdokimos Konstantinidis
Aristotle University of Thessaloniki, Greece

Evangelos Paraskevopoulos
Aristotle University of Thessaloniki, Greece

Antonis S. Billis
Aristotle University of Thessaloniki, Greece

Panagiotis Emmanouil Kartsidis
Aristotle University of Thessaloniki, Greece

Panagiotis Antoniou
Aristotle University of Thessaloniki, Greece

Christos A. Frantzidıs
Aristotle University of Thessaloniki, Greece

Panagiotis Bamidis
Aristotle University of Thessaloniki, Greece

ABSTRACT

Early detection and prediction of dementia through unobtrusive techniques or obtrusive tests is still in exploratory status and despite the increase of interest in recent years, many challenges remain open in designing methodologies that can accurately predict its onset. This chapter addresses the problem of the early detection of dementia from two points of view: Detection based on unobtrusive paradigms both in lab and home environments (behavioral monitoring, serious games, home based assisted living applications in telemedicine) and detection based on neuroimaging approaches. The chapter also provides information on setting up ecologically valid home labs for dementia related experiments. Consequently, the aim of this chapter is to provide an overview of a complete methodology of how researchers can possibly detect or predict the onset of dementia through the current state-of-the-art, underline open challenges and illustrate future work in the field.

DOI: 10.4018/978-1-7998-1204-3.ch098

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Chronic illnesses impose a great burden on the lives of citizens worldwide. In modern health-care, decentralisation, de-hospitalization and self-management of diseases at home are crucial factors for improving the every-day life of the patients and the people close to them. People in general tend to dislike obtrusive monitoring on their daily activities, so the challenge for home care solutions is to implement systems that provide clinicians with adequate and concise information on their patients' health status while at the same time be unobtrusive and easy to use (Golle & Partridge, 2009; Röcker, Hinske, & Magerkurth, 2007). Based on sets of multimodal monitoring parameters and measuring scales feeding reconfigurable event detection mechanisms used for risk assessment and analysis, AAL systems aim to early detect symptoms that predict decline, avoid emergencies and secondary effects and, ultimately, prolong the period that patients can remain safely cared at home. In this chapter the authors aim at providing a description of current trends in context-aware ambient monitoring for the early detection of dementia and prediction of deterioration, as well as provide a background for minimally obtrusive neuroimaging approaches on prediction of dementia.

BACKGROUND

The current applied paradigm for dementia, revolves around patients in the later stages of the disease, where the patient is beyond the stage of the application of delaying/preventive measures. This follows as a result of the delay in seeking care from the majority of patients and caregivers, until the disease has progressed (Boise, Morgan, Kaye, & Camicioli, 1999). There is a gap in the clinical paradigm which does not include the need for the application of methodologies for early detection of dementia, at the onset of the very first symptoms. This is further exacerbated by the fact that most physicians delay providing a diagnosis as they perceive dementia as a "terminal" disease without adequate remedial ways of treatment. These factors put together, severely compromise healthcare quality for such patients and also deprive them of the chance to face their disease at its early stages where much can be done, directly impacting overall patient care.

Our current understanding of dementia and the underlying mechanisms is that the disorder may well begin one or two decades before symptoms appear (Reiman et al., 2012). Dementia itself is identified as a late stage in the pathological development of neurodegenerative disorders. In the recent years, there has been a clear shift towards addressing the disease at its earlier pathological stages, but still the main focus of clinical practice falls on the later stages, with most diagnoses after the moderate or even severe stage. On the other hand, there is a rising trend to turn our focus on earlier detection of dementia, seeking prodromal, MCI or even completely healthy subjects to volunteer at experimental protocols for the early detection of dementia. Here we discern a huge impediment: Although we are in need of patient populations at the very early stages of dementia, to begin constructively developing better and more efficient detection paradigms and thus enabling clinical research in the race for better treatments, most patients are diagnosed very late. These facts may contribute to delayed or undetected cognitive decline or even false diagnosis among primary care providers (Borson et al., 2013; McCormick et al., 1994).

In the race for early detection of dementia, much can be said and more can be assumed or predicted. Following the rising trend of multidisciplinary practices in many scientific fields, thus early detection of dementia is now moving away from pure clinical practice and clinical instruments and moving to-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

wards a holistic approach (Flicker, 1999; Hughes, Robinson, & Volicer, 2005). Making use of machine learning techniques, supporting multimodal data, ranging from biomarkers to behavioral metrics and from social activity to serious games, detection of so far invisible patterns of signs of dementia is now probable, in the near future.

Of course, despite the many advances, there is still a multitude of open problems that need to be addressed. These problems range from the mundane, yet fundamental lack of preclinical or even MCI subjects that are willing to participate in long-term trials, to the organization of experimental trials on home-based or home-like environments for monitoring as well as the more technical need for robust pattern recognition algorithms and decision support mechanisms to fully make use of the not-that-available data.

In 2015, there are over 34 million later stage patients worldwide and the current paradigm must undergo a severe shift towards another direction. By the time of the diagnosis, subjects have mostly declined to the point where they lack the mental capacity to play a role in their care and most certainly are beyond the point of providing useful data to the research community. Moreover, over 50% of patients with dementia go undiagnosed by their primary care provider (Boise et al., 1999; Bradford, Kunik, Schulz, Williams, & Singh, 2009; Kotagal et al., 2015).

Dementia care should and must shift to a different paradigm where focus is more on preventive and early detection methodologies than just remedial action after the onset of the later stages.

On the other hand, recent developments in connectivity analysis for brain functionality through neuroimaging, shows increased potential for early detection of dementia both in clinical environment and in "home-lab" environments –as addition to an AAL platform- through the use of low-obtrusiveness sensors (such as Emotiv)

PATIENT MONITORING ENVIRONMENTS

Before examining the actual details and challenges the problem of early detection of dementia induces, we must initially lift our point of view one level and examine the setup of patient monitoring environments. We have already described how the current established paradigm focuses on monitoring patients after the onset of the disease. Unobtrusive monitoring methodologies are effective when subjects are already in an advanced stage of the disease, but monitoring of healthy -or even MCI subjects- is hindered by the psychological factor: Even though monitoring is unobtrusive, the subject knows that it is being monitored. So, in conjunction with the possible unfamiliar environment (i.e if tests are being performed in the premises of the clinic), we face the problem of changed behavior. The subject is moved away from the familiar comfort zone of the home and being subjected to various tests by clinical staff and strange-looking equipment (McKanna, Jimison, & Pavel, 2009). That of course raises stress levels and leads to perhaps exaggerated reactions to the monitoring protocol. The subject's perception of the monitoring environment directly affects the monitoring parameters (especially behavioral ones and even some physical) so to facilitate valid results we must induce the concept of *ecological validity* in our monitoring environment.

By "ecological validity" we mean that the methods, materials and setting of the experiment must approximate the real-world situation that is being examined (Brewer, 2014). This is particularly important in studies relating to fields such as cognitive psychology, neuropsychology and apply to aspects relating to decision making and judgment, factors of grave importance to the -early or not- detection of dementia and its results on the subjects' behavior.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Apart from ecological validity as a crucial factor in conducting research on the early detection of dementia, we should shortly digress on another important issue, that of the lack of existing data pool. Until the current paradigm shifts towards early detection of dementia, which means that more researchers and more institutions focus their efforts on this, current research lacks from available data on which algorithms, methodologies and detection pipelines can be applied. This combined with the wide timeframe needed (in the order of magnitude of years) to generate meaningful results, leads us to conclude that as much data can be available to as many researchers as possible. The scientific community has to simultaneously produce and share data as widely as possible to facilitate swift progress on the early detection of dementia. Thus, we come to the combined need for ecologically valid monitoring environments built on an open data scheme of dissemination for more accurate data combined with real-time dissemination capabilities.

THE OPEN DATA ECOLOGICALLY VALID LIVING LAB PARADIGM

There are quite a few ongoing research initiatives focused on the creation of integrated unobtrusive smart lab/home environments supporting the independent living of elderly people. The main goal of these smart environments is the ambient monitoring of the elderly through a set of monitoring devices (which comprises a common monitoring model in research efforts across the world). The subjects are continuously evaluated in ecologically valid living lab/home environments and monitoring parameters -behavior, gait, sentiment, activity, memory test results etc- are being recorded and stored -anonymized- for processing and pattern recognition methodologies to be applied on them . Subjects are being analyzed as unobtrusively as possible, just by performing routine tasks they would anyway perform at their homes, or on another level being prompted to perform some light physical exercises, in an overall wellness scheme, through the use of serious games promoting physical exercise or activity. Thus, we collect and can analyze a multitude of data, as little as possible unaffected by perception and stress related distortion.

Such an example of an ecologically valid living lab setup as in Figure 1 would consist of a living-room space, a bathroom-like space and a hall-kitchen space, all in the same room. In this scenario, a number of elderly volunteers visit this environment daily and perform a set of routine activities just as they would at home. An ICT infrastructure (for example, NoSQL databases, Kinect streaming services, etc.) supports the recording of the raw information of all the sources, comprising Kinect(RGB, skeleton and microphone array information), Smart Watch (acceleration), Wii Balance Board (center of gravity during ExerGames interaction, related to the balance ability of the elderly) and health measurements. The data generated by such a monitoring scheme would reach a mean throughput of 10GB per hour swiftly generating a big database of ready to be reproduced, shared and analyzed by the wide scientific community (playback) (Evdokimos I. Konstantinidis, Antoniou, Bamparopoulos, & Bamidis, 2014). On top of this, the paradigm exploits an RDF database contains semantic annotation of the datasets which can be provided by an external observer. In such a monitoring scheme, the infrastructure and ICT tools (streaming on demand, cloud storage and processing, etc.), are based on already developed services and frameworks (such as the cac-framework.com) (Evdokimos I. Konstantinidis et al., 2014) to make these data available on demand. As stated by Konstantinidis et al. (2014), the scientific community will be able to experiment and to test their algorithms on data stemmed from an ecologically valid environment in lab settings. This approach will enable developers and researchers of early detection algorithms that rely on information similar to the existing datasets, by enabling the design and developed algorithms

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. An ecologically valid living lab in laboratory settings resembling a real common living room and kitchen

and software tailored to the elder's daily living patterns. Most importantly a preliminary evaluation of the service adoption has already been conducted on the programmers/developers front with encouraging results.(E. Konstantinidis, Antoniou, et al., 2014; E. Konstantinidis, Billis, et al., 2014).Therefore, such an approach would facilitate the ulterior aim of speeding up innovation, collaboration as well as development and testing of more accurate services, which is achieved by an early involvement of users (playback already existing datasets of elderly in an ecologically valid living lab) as co-creators.

We have stressed the need to produce ecologically valid environment datasets in a real time format, which will support and accelerate the collaboration among the scientific community, as well as in a datasets format for further exploitation of other entities, in a wide-open data scheme that will return back to the research community a wealth of information along with convenient tools for getting access to them.

ECOLOGICALLY VALID DATA COLLECTION

In the current monitoring paradigm, the assessment of a subject's cognitive and functional status is based on brief clinical visits, where various data is collected via short tests and compared with results from previous sessions. This is conventionally done by using a variety of validated clinical assessment tools and is heavily affected by the distorting lens of the subjects' and caregivers' recollection of events in the days preceding the tests as well as their capacity for self-assessment, which is -as already stated- affected by the "sterile" clinical environment. Moreover, infrequency and the actual short duration provide more confounding variables for the accuracy and efficacy of these tests.(Boise et al., 1999; Bradford et al., 2009)

So, the biggest challenge concerning not only the early detection of dementia, but also the more conventional monitoring of the progress of the disease itself, is the general lack of pervasive and unobtrusive environments that will exploit significant parameters, enhance adherence, provide constant monitoring by tracking activities and behaviors and thus a continuum of useful data, with underlying patterns that can be used to predict outcomes of mood, behavior, cognitive status and functionality, used as biomarkers for detecting trends towards dementia (Lyons et al., 2015). Moreover, although there is a

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

plethora of ambient intelligence examples of home monitoring applications, these are mainly focused on life threatening acute events and less on slow varying trends that can indicate future conditions such as dementia (Mihailidis, 2008). Therefore, aiming to fill the above gap, we must focus on utilizing low-cost, off-the-self, unobtrusive sensors and devices to detect both short-term monitoring - such as the detection of emergency events as well as the detection and assessment of ADLs and other contextual situations (Storf, Becker, & Riedl, 2009)– and, long-term monitoring to detect abnormal patterns of health or behavior and continuous health assessment (Chiriac, Röll, Parada, & Rosales, 2012; Monekosso & Remagnino, 2010; Xefteris et al., 2011). Since cognitive impairment is coupled with functional decline and depressive symptoms, (Mehta, Yaffe, & Covinsky, 2002) the need for a multimodal fusion scheme that monitors multiple aspects of the subjects' life is apparent. In this context, we must also keep our mind on the tradeoff between unobtrusiveness and accuracy, coupled with the subjects' adherence. Unobtrusiveness seems to be a critical factor that enhances acceptance and compliance, while on the other hand obtrusive or even invasive methodologies provide more accurate data.

The transition from monitoring the disease to actually detecting it at the very early stages or even years before, calls for data acquisition and processing platforms, collecting and fusing data in a multimodal environment (sensors, exergames, questionnaires, etc.). The trending paradigm's objective is to integrate and extend state-of-the art technologies for:

- Fusing and interpreting the parameters uni-modally extracted from raw sensor data into a semantic representation of human behaviour and physiological status.
- Extracting correlations from sets of fused data (cognitive, emotional and physiological measurements). Utilisation of various soft computing, probabilistic and data mining techniques is crucial in order to operate on sensor data to provide least error prone analysis and decision support indicators.

The main parameters, based on Activities of Daily Living (ADLs) that can provide us with useful data, collected unobtrusively, for the possible early detection of dementia and help us make predictions are analysed in the following paragraphs.

Mobility

Since 1998 (Camicioli, Howieson, Oken, Sexton, & Kaye, 1998) has demonstrated that reduced mobility precedes cognitive impairment in the elderly. Researchers now, develop new algorithms for continuous in-home assessment of both variability in movement speeds and movement patterns, to discern underlying signs of risk. On one hand, researchers can employ a scheme of "passive" monitoring via sensors installed in various places inside a home and collecting data on speed, frequency of passage, total activity etc, or by "active" monitoring via the application of ExerGames that both aim to promote and sustain a healthy and active lifestyle and collect possible "anomalous" data in the performance of the various exercises.

Passive monitoring via dispersed sensors has many advantages but also some caveats. On one hand, passive, unobtrusive sensors solve the problem of compliance and adherence, because monitored people tend to either avoid or even forget to wear the common actigraphy devices used for mobility monitoring. Moreover, simple accelerometers do not provide location information - which is significant itself- coupled with the fact that since in-home movement is generally slow and non-linear, data interpretation is very difficult. Thus, in the case of older adults, the notion of ecologically valid unobtrusive monitoring comes

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

back again as a very important factor. To that end, arrays of in-home sensors can be deployed supporting information collection of walking speed, location and frequency of passage. (Hagler, Austin, Hayes, Kaye, & Pavel, 2010; Hayes et al., 2008). The challenge remains though, to establish the individual activity patterns and manage to measure meaningful change that may have applications for predicting health parameters and functional status (Hagler et al., 2010) .

Exergames

The emergence of modern interaction interfaces (tablet, laptop with touch screen, smart tv) has been followed by the trend of transferring paper and pencil screening tests to computerized forms which are more acceptable from elders (Collerton et al., 2007). On top of this, and aligned with Plato's statement that "...you can discover more about a person in an hour of play than in a year of conversation..."(Robert et al., 2014), these computerized tests started turning into games with assessment aspects. Introduced with the term "Serious Games", these games are considered to provide more reliable assessment information compared to conventional methods (Bellotti, Kapralos, Lee, Moreno-Ger, & Berta, 2013; Robert et al., 2014). This can be attributed to the fact that users do not perceive serious games as a stressful testing procedure (Cassady & Johnson, 2002).

The assessment aspect of the cognitive serious games contribute to the early detection of neurological disease (Hagler, Jimison, & Pavel, 2014), while exergames, serious games promoting physical activity and exercise (S. Yang, 2010), are deemed as measuring physical health unobtrusively tools (E. Konstantinidis, Billis, et al., 2014; Staiano & Calvert, 2011). This has recently started to change and exergames have just introduced in the field of cognitive assessment(E. I. Konstantinidis, Bamidis, Billis, Kartsidis, & Papageorgiou, 2015). Only recent studies concerned correlations of physical activity features of SGs with cognitive assessment utilizing virtual environments and treadmills and correlating muscle strength and walking time with global cognitive measurements (Tarnanas et al., 2013). Touch screen serious games and a diversity of serious game have been lately employed towards cognitive Parkinson's Disease prediction(S. Liu, Shen, Mei, & Ji, 2013) and cognitive performance inference (Byun & Park, 2011) respectively. Konstantinidis et al. (2014) introduced the FitForAll exergaming platform towards physical exercise promotion. However, its potential role as contributor to the detection of early cognitive decline signs classifies it to a distinct position. The in-game metrics analysis exhibits classification accuracy greater than 72% when discriminating cognitively normal from mild cognitive impaired elderly. This study show also that FitForAll was more accurate than MMSE and comparable to the MOCA.

Sleep

Sleep patterns comprise a crucial indicator of various health problems, both standalone and as associated symptoms (i.e of pulmonary and cardio-vascular problems). In addition, sleep disruption have been identified as correlated with MCI and Alzheimer's disease. (Lyketsos et al., 2002; Tractenberg, Singer, & Kaye, 2005). Lately, sleep disruption has been linked with decreased performance on memory tests even on cognitively intact elderly (Seelye et al., 2015). In this context, it is obvious that in the framework of an unobtrusive monitoring platform that focuses on the early detection of increased risk factors for dementia, sleep patterns provide rich background information. Making use of off-the-shelf sensors like pressure mats and wrist-worn actigraphy monitors, we can employ a scenario of long term monitoring and assessment spanning months rather than a single day, providing a robust baseline that

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

excludes outliers that skew our data and provide comparison backgrounds for between healthy and MCI subjects. In such a study (Hayes, Riley, Mattek, Pavel, & Kaye, 2014)it was found that amnestic MCI (aMCI) patients typically exhibit less sleep disruptions than non-amnestic MCI (naMCI) patients, who in their turn exhibited less than healthy (cognitively intact) individuals. Seelye et al.(2015) also showed that sustained periods of sleep disruptions have a clear impact on cognitive function in healthy adults. Duration of sleep was not related to the cognitive test performance. These findings suggest that continuous assessment of sleep patterns does provide significant information on factors that can contribute to variable or decreased cognitive function.

Social Engagement

Social engagement is an indirect sign of cognitive status. Moreover, higher levels of social activity have been shown to be a preventive measure for the avoidance of cognitive decline. (Akbaraly et al., 2009; Fratiglioni, Wang, Ericsson, Maytan, & Winblad, 2000; Scarmeas, Levy, Tang, Manly, & Stern, 2001)

As researchers now have a better grasp of the relationship between social engagement and cognitive decline, we can use this information in a twofold way:

1. Deployment of methodologies for the early identification of cognitive decline risks and
2. Development of intervention frameworks based on social engagement as a tool to improve cognitive function.

The greatest advantage of using social engagement metrics as cognitive decline risk for individuals is that most of its aspects can be easily measured in a passive way. These can include time spent out of home, frequency and duration of telephone use, frequency and duration of visitors as well as internet use (e-mail, chat etc).

Decision Support

In the previous paragraphs we examined the milieu of a patient monitoring paradigm that is established in an ecologically valid environment, with open access to monitoring data and input parameters spanning activities of daily living (ADL), behavioral aspects and physiological measurements with the express goal to detect adverse signs in order to enable the early detection of dementia. In this paragraph we go behind the monitoring front-end and try to provide a brief working example of a context-aware system that processes this data via pattern recognition and decision support methodologies to provide detection and prediction results on the subject's status. (Xefteris, Andronikou, Tserpes, & Varvarigou, 2010)

We will use the example of the USEFIL project (Artikis et al., 2012; Billis et al., 2015) which provides both short term monitoring for the detection of emergencies through context-aware detection and assessment of ADLs (Storf et al., 2009), and long-term monitoring for the detection of baseline divergent (abnormal) patterns of health, behavior and general health assessment (Chiriac, Röll, Parada, & Rosales, n.d.,2012).

USEFIL is an innovative data acquisition and processing platform that collects and fuses multimodal data (from sensors, exergames, questionnaires etc). Raw data collection is performed in-home and initial data are processed locally to provide user reminders and alerts, with group-level processing performed on a central database.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The DSS component of the USEFIL platform endeavors to recognize early deterioration signs for a multitude of medical problems as well as deviations from a behavioral baseline that indicate decline, alerting both formal and informal carers as soon as possible. The component performs these operations with the use of data mining and pattern recognition techniques in three steps:

1. Collect of the relevant data that is already filtered and pre-processed.
2. Project data segments into a time-series to enable pattern recognition.
3. Classify the occurring temporal fluctuations to diagnosis types.

The USEFIL DSS system is thus designed as a spatio-temporal model comprising A.I methods that process data related to user activity and physiological classification. The DSS utilizes soft computing, probabilistic and data mining techniques on the sensor acquired data to perform least error prone analysis and decision support.

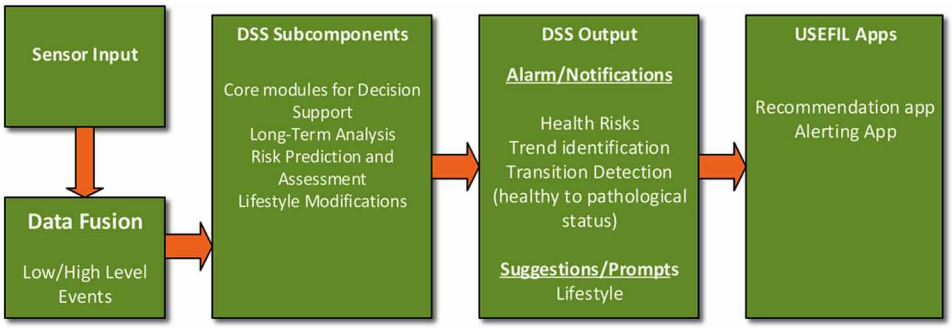
Aiming to accomplish the aforementioned objectives, the USEFIL DSS system is designed to be a spatio–temporal model consisted of artificial intelligence methods manipulating data related with user activity and physiological classification. The progress beyond the current state-of-the art of the specific component is:

- Both short- and long-term monitoring.
- Hybrid decision support system, based on rule-based expert systems combined with fuzzy cognitive maps (FCM) -a type of fuzzy inference scheme.
- Involvement of medical experts in the design process of the DSS, in order to improve the targeting and acceptance of the system with professional carers (Papageorgiou, Billis, Frantzidis, Konstantinidis, & Bamidis, 2013).

The conceptual approach of the DSS component is shown in Figure 2.

The concept followed is that within the Data Fusion Tool the raw data are fused and activity states or transitions among them concerning behavioral and physiological states are stored as events. (Katzouris, Artikis, & Paliouras, 2014)

Figure 2. USEFIL DSS conceptual architecture (I/Os and modules)



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The Sensor Data Fusion tool outputs these events to the other DSS subcomponents that take care of detecting the various possible pathological states, extrapolate the possibility for future deterioration and identify possible pathological trends as well as suggest remedial approaches. These outputs are categorized in alarms/warnings/notifications or into prompts and suggestions, providing triggering flags to other USEFIL applications (such as an Alerting App) that interact with the subject or the carer. Long term analysis results may be also used as input to the other components (Billis, Papageorgiou, Frantzidis, Konstantinidis, & Bamidis, 2013).

The flowchart in Figure 3 provides a depiction of how the different DSS components are interconnected. (Panagiotis D. Bamidis, Tarnanas, Tsolaki, & Hadjileontiadis, 2015). The various outputs can be considered as standalone DSS results or be used by other DSS components for the production of further conclusions. As an example, the presence of severe symptoms can be used as a standalone result of the long-term analysis component or be combined with the DSS to classify a senior as suffering from a specific pathological condition.

These detected trends may be also used by the Risk Assessment Component, in order to provide estimation of potential future adverse situations. The DSS forms a "pool of evidence" comprising estimation of health trends, current health status and possible future outcomes, in order to procure the appropriate strategies for the alleviation of health problems and for helping the subject stay independent. The produced strategies are proposed via the "Lifestyle Modifications Suggestion". This transforms the DSS from a simple passive monitoring application to an active suggestion system, not only detecting and predicting, but also being actively engaged in possible remedial actions.

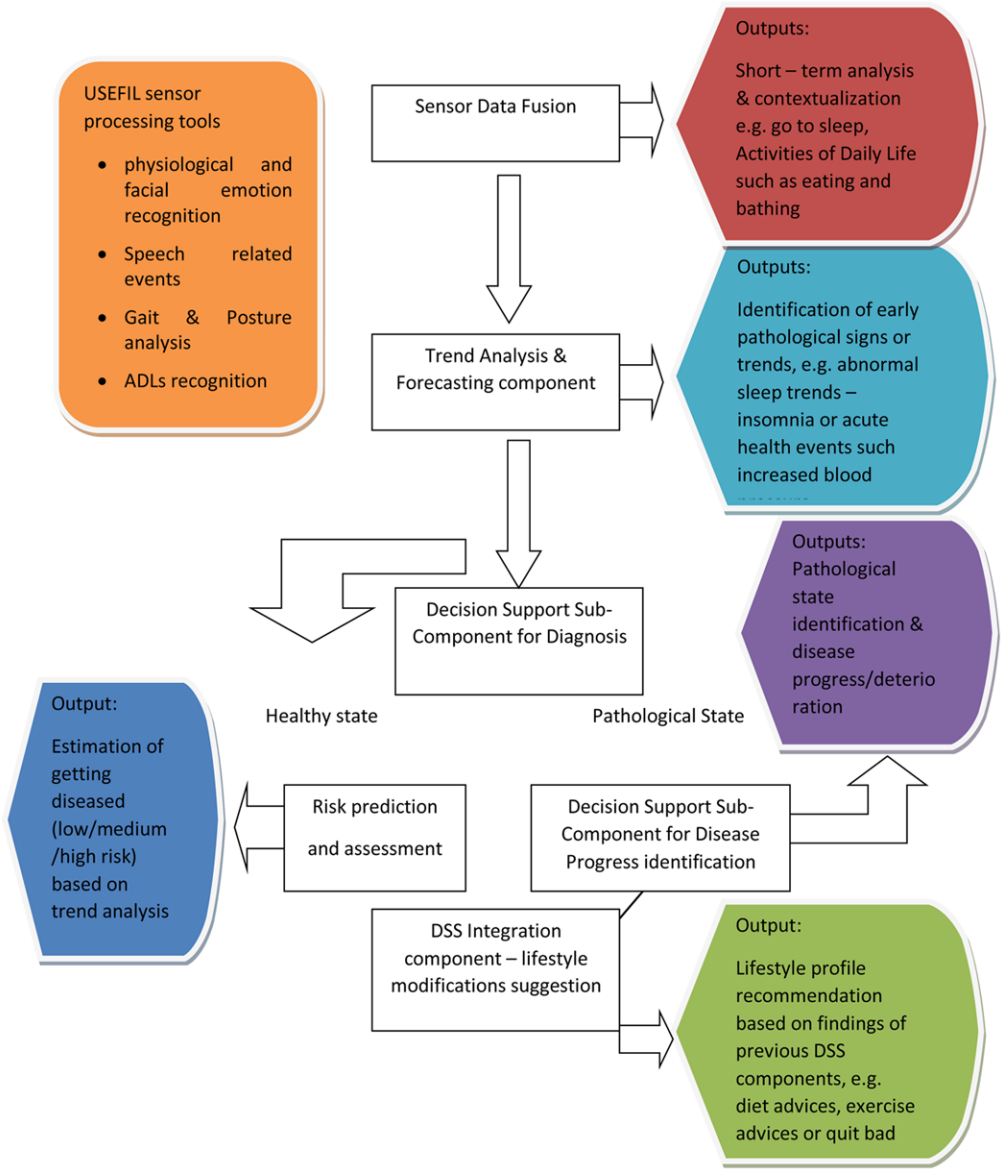
The USEFIL DSS approach is encapsulating the essence of the early detection paradigm by actively engaging in long term monitoring for the identification of adverse trends, providing significant forecasting capabilities for the prognosis of potential future deterioration while at the same time providing possible solutions for delaying this deterioration by the application of certain life-style modifications.

In Vitro Detection with Neuroimaging

We have so far discussed the efficacy of unobtrusive methodologies as reliable markers for the detection of dementia. Nevertheless, reliable biomarkers for the early detection of dementia can be obtained from obtrusive methodologies as well, namely the most common techniques (see next section) usually employed in the field of neuroimaging. Indeed, recent research has shown promising results in the detection process by adding distinctive and reliable neuroimaging biomarkers to the diagnostic criteria of Alzheimer's disease (AD). For instance, (Dubois et al., 2007) demonstrated a significant potential for methodologies that aim to keep cognitive functions active or enhance their performance across the adult lifespan. Engaging seniors into computer based interventions of training schemes tailored to their specific capacities shows great promise in allowing for the longitudinal evaluation and monitor of their cognitive status (Bamidis et al., 2014; Busse, Gil, Santarém, & Filho, 2009; Colcombe & Kramer, 2003; Tardif & Simard, 2011). In parallel, our current understanding of neuroplasticity and the employment of neuroimaging biomarkers have the potential to offer additional solutions in the existing clinical routine for early detection and also for monitoring the disease's progression. The aim of this paragraph is to present recent neuroimaging findings on the dementia detection that can be employed to unravel the functional and structural reorganization of the aging brain.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The interconnections between the USEFIL DSS components



Neuroimaging Modalities

The process of aging induces changes in the brain and directly in the activity of neurons and subsequently, the cognitive function. With recent advances and techniques, we are now able to study both brain structures and functions. Though obtrusive in nature, techniques such as EEG, MRI, fMRI, MEG, and DTI are non-invasive, painless and easy to apply, thus quite significant in the study of the aging brain in healthy as well as pathologic subjects. These methods enable us to reconstruct images of the structural (grey and white matter) and/or functional (electrophysiological, blood flow, metabolic) characteristics

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

of the brain, in spatial scales ranging from a millimetre to several centimetre and temporal scales from a millisecond to hours. Each methodology has strengths and caveats in terms of resolution and thus, the decision on which to apply is dependent on the application.

Brain Plasticity

The process of aging is considered to be a progressive state of continuous decline. New research findings though provide us with quite a different milieu: It is apparent now that the brain retains its adaptability and flexibility across a person's lifespan and many cognitive abilities are preserved further than previously thought.

Plasticity refers to the ability of the brain to reorganize both structural and functional components. Plasticity is an ability that in healthy subjects remains active across the lifespan (Boyke, Driemeyer, Gaser, Buchel, & May, 2008).In contrast to an earlier notion that age-related neuroplastic changes were the result of cell death or dendritic deterioration, it is now well accepted that they are the result of lower synaptic densities in older adults (Hedden & Gabrieli, 2004). The aging process alters the biophysical properties of neurons and this affects the brain network dynamics resulting in poor signal to noise ration (SNR) (Burke & Barnes, 2006)and lower neuromodulatory system function in older brains (Mahncke et al., 2006). These changes mainly affect the prefrontal cortex and the medial temporal structures.

Therefore, intensive training of cognitive or motor processes, can stimulate brain plasticity, inducing changes in the underlying cortical structures (Mahncke et al., 2006). For instance, recent MEG findings indicate that two weeks of intensive multisensory training may induce neuroplastic changes in the functionality of frontotemporal cortical regions in adults (Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2014). Inducing behavioral interventions enables the ability of the brain to reorganize itself, even in the case of the aging (Mahncke et al., 2006)

Since the cognitive deterioration of the brain may onset well before the aging process begins, and changes in plasticity at brain synapses combined with the loss of neurons are the key indicator for oncoming functional decline, it is apparent that detecting these changes at an early stage, can be instrumental in the early detection of dementia. These changes are precursors to dementia related impairments and precursors of AD, such as amyloid beta deposits, brain inflammation, tau tangles and ApoE4 proteins. They directly affect areas of the brain responsible for learning, memory and cognition by impairing synapses and contributing to cell death. Early neuron loss and changes in synapse function have been observed in the hippocampus and the neocortex, the regions involved in language, memory and higher cognitive functions, so using neuroimaging to detect such changes can alert clinicians to the future onset of dementia and provide a route to preventive actions, rather than the established remedial paradigm.

Structural Neuroimaging in Aging and Dementia

Age related structural changes in the brain have varying extents. The threshold between healthy and pathological aging is very thin as many of these brain changes extend in dementia and AD. The aging of the brain can range from a healthy (successful) state, where structure and volume is completely retained until almost the end of our lives, to a more typical or near pathological aging. Some typical dementia related structural changes can be also a part of a healthy aging curve in middle aged subjects as well as elderly subjects who are essentially cognitively intact. Some extent of brain atrophy can be present in healthy aging but when followed by cognitive decline it may induce an increased risk of neurodegenerative

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

diseases. In this context, structural neuroimaging can be instrumental in detecting preclinical stages of age-related cognitive decline and be of great help in the early detection of the progression of dementia. In order to efficiently characterize age-related disorders such as dementia, we need to disentangle the mechanisms of the structural changes (i.e. atrophy) that take place in healthy aging.

MRI Detection of Dementia

In current clinical practice, structural MR imaging is widely available and reasonably inexpensive. The main challenge of MRI, concerning early detection of dementia or other neurodegenerative diseases is to discern MCI patients that will in the future fulfil the clinical criteria for dementia or AD. MRI can also be used to distinguish between non-AD and AD dementias -for instance, hippocampal atrophy is absent in dementia with Lewy bodies (McKeith et al., 2005)and vascular dementia (Barber et al., 1999). MRI is an efficient tool for the very early diagnosis of an individual suspected of dementia and its subtyping, through detection of hippocampal reductions. (Shi, Liu, Zhou, Yu, & Jiang, 2009). Moreover, MRI is employed for predicting the risk of AD progression in populations that are both cognitively healthy and MCI.

Though more research is needed to establish the clinical value of MRI in MCI, measuring these significant reductions due to AD can have a powerful impact on the early diagnosis of MCI and AD. In addition, MRI is employed for the prediction of the risk of dementia progression. (Risacher et al., 2009) established that the MRI biomarker measuring the degree of neurodegeneration of MTL structures was the best for imminent conversion and the biomarker of decreased hippocampal volume the most robust.

So, having established the as early as possible detection of the disease, MRI metrics such as brain volume fluctuations through time, can be used to evaluate and keep track of the disease (Fox & Freeborough, 1997).

DTI Detection of Dementia

Changes in grey and white matter are promising biomarkers of dementia. These markers are detected by microstructural abnormalities on DTI-based diffusivity indexes. These alterations enable the prediction of dementia progression independently of possible volume loss in prodromal situations (Canu et al., 2010; Fellgiebel et al., 2006; Oishi, Mielke, Albert, Lyketsos, & Mori, 2012) and are evident in both healthy brain aging as well as in MCI and dementia, but also in other neurodegenerative disorders (Kantarci et al., 2011; Oishi et al., 2012; Rogalski et al., 2012) .

In MCI patients, higher baseline diffusivity in the hippocampal head indexes a bigger risk of progression to AD. Thus, diffusion findings can be employed as complementary biomarkers to hippocampal atrophy in identifying progression. When taken in earlier stages of life, these findings can be indicators of the future appearance of symptoms.

Functional Neuroimaging and Resting State Analysis

To understand the effects of aging on neural plasticity, we have to initially examine the effect of age on the brain's functionality. Functional neuroimaging detects age-related increases and decreases of neural activity, in conjunction with cognitive decline. These approaches provide a link between the various implications of aging on our brains by measuring either resting state activity or neural activity while subjects perform cognitive tasks.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In this context, neuroimaging studies make use of resting state measurements for the definition of a baseline which is later compared with the results from the performed tasks. At rest, the cortex's functional characteristics show a high degree of reproducibility and stability and some brain regions show a consistent activity than during cognitive tasks (Greicius, Srivastava, Reiss, & Menon, 2004). The implicated brain regions include the posterior cingulate cortex (PCC), MTL structures, dorsal and ventromedical PFC and lateral (mainly inferior) parietal cortices (Mevel, Chételat, Eustache, & Desgranges, 2011). Due to its functional characteristics this is dubbed the Default Mode Network (DMN) (Raichle et al., 2001) . The DMN is characterised by a highly stable functionality which makes it a valuable tool for the evaluation of cognitive and neurological disorders such as amnsestic MCI (aMCI) (Garcés et al., 2014) and AD (Greicius et al., 2004).

Using fMRI for Resting State Analysis

fMRI is a highly promising biomarker for dementia, repeatable many times over the course of longitudinal studies. fMRI studies examining the effects of dementia in activation patterns and organization of DMN of dementia patients (Y. Liu et al., 2008) and age-matched healthy subjects revealed that dementia groups exhibit decreased resting state activity in almost all of the cortical regions involved in the DMN and most prominently in the hippocampus and the PCC, thus suggesting disrupted connectivity between these two regions in dementia(Greicius et al., 2004). In an earlier study (Shi et al., 2009)it was established that dementia patients exhibit diminished connectivity of the hippocampus to other regions, compared to healthy elderly individuals. In this study, a group of MCI patients had intermediate values, also significantly different from the dementia group. Lack of presence of this connectivity breakdown in the visual cortex suggests that the sensory systems may be affected only in later stages of the disease. Studies also revealed that the functional connectivity of the hippocampus with the medial PFC, ventral anterior cingulate cortec (ACC), right inferior temporal cortex, right cuneus (Cu) and precuneus (PCu), left Cu, right superior and middle temporal gurys and PCC was disrupted in dementia patients compared to healthy adults (L. Wang et al., 2006)

In Sheline et al.(2010) it is revealed that the functional connectivity breakdown of the DMN that dementia patients exhibit might predate the pathological effects of amyloid plaque toxicity.

Thus, detecting early connectivity breakdowns in the DMN via fMRI in resting state, could be of big help in a future early dementia detection paradigm.

Using MEG-EEG for Resting State Analysis

EEG has a high temporal resolution which makes it an ideal tool for studying the cortical connectivity of DMN. Several studies employed this approach searching for disrupted connectivity between cortical regions of the DMN, indicative of neurodegenerative diseases.

In a 2014 study, (Frantzidis et al., 2014) used EEG to reveal a significant reduction in the small-world brain architecture in both aMCI and mild demented patients relative to the group of healthy controls. These findings correlated significantly with the cognitive function as evaluated by MMSE. All of the above are in line with the notion that the AD may be grounded on a connectivity breakdown and may actually constitute a disconnection syndrome, as suggested by Delbeuck (Delbeuck, Collette, & Van der Linden, 2007).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Apart from structural and connectivity analyses, several studies have used another approach and ventured to identify the electrophysiological effects of dementia and MCI to the functionality of the DMN. These studies are focused on comparing the EEG activity measured across different frequency bands between dementia, MCI and healthy subjects. The results of such studies indicate that dementia patients in contrast to healthy subjects exhibit increases in the power of low frequency bands such as delta (-4Hz) and theta (4-7 Hz) in central and frontal electrode sites as well as a decrease in the power of posterior alpha (8-15Hz) and beta (16-30Hz) (Babiloni et al., 2004, 2013; Bennys, Rondouin, Vergnes, & Touchon, 2001). These findings are corroborated by MEG studies which exhibit increased power in the activity of parietal delta sources(Fernández et al., 2006) and more significantly, these abnormalities seem heavily correlated with global impairment of cognitive function via an MMSE evaluation (Jeong, 2004).

Comparing MCI patients with healthy controls, reveals the same pattern of findings (increase in power of low frequency bands in posterior sites and decrease of alpha and beta bands for MCI patients). Thus, we are led to pose the hypothesis that there can be a linear transition point between healthy and MCI and MCI and dementia, and that these EEG findings may provide stable screening biomarkers for the risk of onset of dementia and prediction regarding the progression from MCI to AD(Babiloni et al., 2013). Recently, an EEG study explored the effects of combined cognitive and physical training on resting state brain activity of MCI patients in comparison to the effects of solely cognitive or physical training as well as in comparison to active and passive control groups. The effect of the combined intervention was interpreted as beneficiary neuroplasticity changes (i.e. decrease delta and theta rhythms in the precuneus/posterior cingulate cortex) across MCI patients and these changes correlated with cognitive status improvement.(Styliadis, Kartsidis, Paraskevopoulos, Ioannides, & Bamidis, 2015).

FUTURE CHALLENGES

In 2013, over 35 million people around the world were estimated to have dementia.(Camicioli et al., 1998) The number of people worldwide with dementia is expected to increase to 65.7 million in 2030 and 115.4 million in 2050, which is approximately a doubling of sufferers every 20 years. ("WHO | Dementia: a public health priority," 2012). Dementia is a syndrome of progressive cognitive and functional deterioration which is incurable. There are diverse symptoms and various levels of impairment depended on affected brain regions and types of dementia. Contemporary medical technologies related to dementia have proven track record for detection and intervention but have also been distilled in more mild modalities to be implemented in the case of less severe needs. Recent technological advances that allow the in-vivo study of brain structures and functions in human subjects have greatly contributed to a better understanding of the brain changes that occur due to aging and are likely to impact the cognitive function and neuronal activity. The non-invasive nature of most of these techniques (i.e. Electroencephalography – EEG; Magnetoencephalography – MEG; Magnetic Resonance Imaging – MRI; functional MRI – fMRI; Diffusion Tensor Imaging – DTI) makes them an invaluable tool in the study of the living brain in healthy as well as pathologic subjects (Styliadis, Kartsidis, & Paraskevopoulos, n.d.).

With diagnostic modalities providing significantly early detection of various dementia types the question becomes one of choice of interventions. Additionally, with studies outlining depressive symptoms, as a consistent factor which is negatively related to self-rated QoL of people with dementia living in long-term care facilities (Beerens, Zwakhalen, Verbeek, Ruwaard, & Hamers, 2013) as well as the self-evident costs related with long term care the Ambient Assisted Living (AAL) approach to the

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

maintenance of quality of life and health care interventions provide greater flexibility and respectful interventions at the behavioral level with unobtrusive and transparent monitoring and interventions at familiar home environments.

Beyond the, now established, monitoring with unobtrusive sensors and devices that are natively ICT enabled an emergent field in elderly healthcare is the Internet of things (Ashton, 2009). Transparent and ubiquitous communication capacities by ad-hoc networking of various devices and actuators has begun finding application in elderly healthcare. This technology has been demonstrating, rather recently -e.g. (G. Yang et al., 2014), a significant number of enabling capacities in the field of elderly healthcare, especially for people with cognitive decline. Inkjet printed, flexible bio-patches that allow for portable and very unobtrusive ECG measurements; RFID enabled medication packaging that can sense how many medication slots have been opened, facilitating dramatic increase in medication adherence of elderly people; unique RFID tagged medical ids linked to detailed medical records for protection against false medical history. Such unobtrusive and transparent technologies provide critical quality of life improvement and subsequent lifestyle and healthcare maintenance.

With all these advances in ubiquitous networking and unobtrusive monitoring an area that appears to not have been heavily tapped into is the technologically enabled communication and collaboration between elderly that are suffering from one or another form of non-trivial age based cognitive decline. The effect of an existing, active and supporting social network in real life has been explored providing evidence that it is associated with the well-being of the elderly (X. Wang, 2014). What hasn't been explored so far in sufficient detail is the extent in which an electronic social network of dementia patients and their carers would impact in their quality of life. Although folksonomies (folk-generated taxonomies from custom subject tagging in social networks and forums (Adamic & Adar, 2005)) have led to emergent social networking the use of such provisions are still not sufficiently explored. In the context, however, of AAL environments and protected but ubiquitous networking there is significant potential for unobtrusive behavioral monitoring and intervention in the, now widely accepted, form of an electronic social network.

From the previous discussion, distilling content from this whole work three overall monitoring and intervention feedback loops emerge. In the purely biomedical level the aforementioned brain activity and anatomy monitoring and imaging capacities (e.g. EEG, PET, fMRI) provide immediate feedback to the medical doctor for medicinal or non-medicinal treatment of dementia's underlying pathogeny at the purely neurophysiological level. On a different level, ambient, unobtrusive sensor based monitoring and assessment in the home environment provides the necessary data that are usable for prescribing and suggesting personal behavioral modification that will improve or at least maintain quality of life (e.g. physical exercise programs, exergaming interventions). Furthermore, even though largely unexplored in the context of ICTs, real world support group interventions of elderly with varying degrees of dementia provide to the expert crucial data for intervening at the social level and creating a fertile social environment for the patients to maintain or even improve their way of life (Leung, Orrell, & Orgeta, 2015).

Medical diagnostic modalities, intervention protocols, unobtrusive ambient home monitoring and intervention as well as social support infrastructure have the capacity of creating an immense amount of data. From patient records and imaging to real time sensor data these are sources of information that can provide a wealth of additional meta-information using context, time evolution and aggregation techniques. In order to achieve "human activity recognition" (Rashidi & Mihailidis, 2013) a number of techniques must be applied in order to transform raw sensor data to human activity patterns. In that endeavor methods of either supervised learning through annotation, using annotated data and supervised machine learning algorithms can be utilized, or, even more abstractly, unsupervised data mining

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

techniques can be incorporated (Rashidi & Mihailidis, 2013). These algorithms, in order to cope with the needs of interpreting activity in various contexts as well as discriminating outliers with significant anomalies in activities have widely adopted the use of modelling metadata and ontologies in order to codify both context and underlying semantic relationships.

Evolving from information theory, efforts are made to explain and exploit the cross-scale information processing that is apparent in natural dynamical systems. By viewing physical, biological, or social systems as networks of information exchange, statistical tools from information theory can be used to causally explain and quantify the effect that each participant of this network contributes to the overall behavior of the network (Quax, Apolloni, & Sloot, 2013). Through such toolbox, outcomes from contemporary physics, such as resonance phenomena and non-trivial propagation phenomena can be quantified and predicted (Czaplicka, Holyst, & Sloot, 2013). In that way specific structural changes in a network's characteristics can dramatically alter the dynamic evolution of various metrics of the system that is described by it. Such scientific paradigms can and have been used to explore the not only the fundamental processes on different and separate scales of dynamic systems, but also the cross-scale coupling of them and any possible resulting emergent properties that can be exploited for pattern recognition or efficient interventions (e.g.("dynanets.org," 2015)).

This kind of analysis tools have found use in a variety of disciplines and fields where implementing a specific hypothesis about a network, based on the characteristics of the participating nodes can lead to tangible conclusions about macroscopic and measurable characteristics of the whole system. For example information dissipation length (IDL) has been introduced as an index of global instability of the time series of USD and EUR interest rate swaps (IRS) presenting a specific pattern of gradual increase and subsequent decrease before and after the Lheman Brothers collapse, respectively (Quax, Kandhai, & Sloot, 2013). On another example by introducing the concept of transmission networks and proposing a filter reduction method in combined data from social networks and phylogenetic analysis, tangible results regarding correlation of cross-group sexual behavior and HIV transmission has been found (Zarrabi et al., 2013).

Bringing all these themes together, smart sensors and semantic technologies provide aggregated high level data about daily activities. Contemporary decision support and machine learning technologies can consume such high level data streams, in real time, providing intervention suggestions at both the physical and mental pathogeny level. Neuroscience advances provide assessment tools able to rapidly validate, and intervene with gamified and collaborative activities for mental health and overall wellbeing. What is at the moment missing is a formalization, both at a high level of abstraction, as well as at the prototype implementation level, of the role of these technologies and interventions, in the paradigm of maintaining wellness through structured interaction. Utilizing the aforementioned information theory background. Such a synergetic use of these technologies would provide both a conceptual paradigm and an implemented assessment and intervention prototype for more effective and impactful dementia monitoring and intervention. From this discussion of the previous sections on modalities and means, combining the previous discourse it becomes apparent that, apart from following through in researching specific aspects of dealing with the dementia patient, a significant future challenge will be the integration of the technologies and methods into something beyond an incoherent mashup of randomly interacting parts. Detailed data aggregation methods can provide detailed but also seamless exploration of various modalities' and interventions' impact. This methodology can explore, utilizing the previous theme of cross scale feedback propagation, the necessary exact prescriptions for all the components of this palette for achieving rigorous integration to protocols. In that effort the support from ecologically

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

valid dementia oriented living labs in conjunction with the treatment space (hospital, personal physician) become an organic research testbed for this integration. Only through such a coordinated infrastructure of integrated research can, large scale trials explore the cross-scale and cross-modal impact of an emerging ecosystem of dementia prevention, treatment and care in order to provide an appropriate, immediate and personalized monitoring and intervention "prescription" that will ensure each individual patients length and quality of life.

REFERENCES

Adamic, L., & Adar, E. (2005). How to search a social network. *Social Networks*, *27*(3), 187–203. doi:10.1016/j.socnet.2005.01.007

Akbaraly, T. N., Portet, F., Fustinoni, S., Dartigues, J.-F., Artero, S., Rouaud, O., ... Berr, C. (2009). Leisure activities and the risk of dementia in the elderly: Results from the Three-City Study. *Neurology*, *73*(11), 854–861. doi:10.1212/WNL.0b013e3181b7849b PMID:19752452

Artikis, A., Bamidis, P. D., Billis, A., Bratsas, C., Frantzidis, C., Karkaletsis, V., … Spyropoulos, C. S. (2012). *Supporting tele-health and AI-based clinical decision making with sensor data fusion and semantic interpretation: The USEFIL case study*. Retrieved July 24, 2015, from http://airtlab.dii.univpm.it/netmed2012/sites/netmed2012.dii.univpm.it/files/papers/netmed2012.pdf#page=32

Ashton, K. (2009). That "Internet of Things" Thing. *RFiD Journal*, *4986*. Retrieved from http://www.itrco.jp/libraries/RFIDjournal-That Internet of Things Thing.pdf\npapers3://publication/uuid/8191C095-0D90-4A17-86B0-550F2F2A6745

Babiloni, C., Binetti, G., Cassetta, E., Cerboneschi, D., Dal Forno, G., Del Percio, C., ... Rossini, P. M. (2004). Mapping distributed sources of cortical rhythms in mild Alzheimer's disease. A multicentric EEG study. *NeuroImage*, *22*(1), 57–67. doi:10.1016/j.neuroimage.2003.09.028 PMID:15109997

Babiloni, C., Carducci, F., Lizio, R., Vecchio, F., Baglieri, A., Bernardini, S., ... Frisoni, G. B. (2013). Resting state cortical electroencephalographic rhythms are related to gray matter volume in subjects with mild cognitive impairment and Alzheimer's disease. *Human Brain Mapping*, *34*(6), 1427–1446. doi:10.1002/hbm.22005 PMID:22331654

Bamidis, P. D., Vivas, A. B., Styliadis, C., Frantzidis, C., Klados, M., Schlee, W., ... Papageorgiou, S. G. (2014). A review of physical and cognitive interventions in aging. *Neuroscience and Biobehavioral Reviews*, *44*, 206–220. doi:10.1016/j.neubiorev.2014.03.019 PMID:24705268

Barber, R., Gholkar, A., Scheltens, P., Ballard, C., McKeith, I. G., & O'Brien, J. T. (1999). Medial temporal lobe atrophy on MRI in dementia with Lewy bodies. *Neurology*, *52*(6), 1153–1153. doi:10.1212/WNL.52.6.1153 PMID:10214736

Beerens, H. C., Zwakhalen, S. M. G., Verbeek, H., Ruwaard, D., & Hamers, J. P. H. (2013). Factors associated with quality of life of people with dementia in long-term care facilities: A systematic review. *International Journal of Nursing Studies*, *50*(9), 1259–1270. doi:10.1016/j.ijnurstu.2013.02.005 PMID:23465959

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of Serious Games: An Overview. *Advances in Human-Computer Interaction*, *2013*, 1–11. doi:10.1155/2013/136864

Bennys, K., Rondouin, G., Vergnes, C., & Touchon, J. (2001). Diagnostic value of quantitative EEG in Alzheimer's disease. *Neurophysiologie Clinique. Clinical Neurophysiology*, *31*(3), 153–160. doi:10.1016/S0987-7053(01)00254-4 PMID:11488226

Billis, A. S., Papageorgiou, E. I., Frantzidis, C., Konstantinidis, E. I., & Bamidis, P. D. (2013). Towards a hierarchically-structured decision support tool for improving seniors' independent living. In *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '13* (pp. 1–4). New York: ACM Press. http://doi.org/10.1145/2504335.2504362

Billis, A. S., Papageorgiou, E. I., Frantzidis, C. A., Tsatali, M. S., Tsolaki, A. C., & Bamidis, P. D. (2015). A Decision-Support Framework for Promoting Independent Living and Ageing Well. *IEEE Journal of Biomedical and Health Informatics*, *19*(1), 199–209. doi:10.1109/JBHI.2014.2336757 PMID:25073180

Boise, L., Morgan, D. L., Kaye, J., & Camicioli, R. (1999). Delays in the diagnosis of dementia: Perspectives of family caregivers. *American Journal of Alzheimer's Disease and Other Dementias*, *14*(1), 20–26. doi:10.1177/153331759901400101

Borson, S., Frank, L., Bayley, P. J., Boustani, M., Dean, M., Lin, P.-J., ... Ashford, J. W. (2013). Improving dementia care: The role of screening and detection of cognitive impairment. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, *9*(2), 151–159. doi:10.1016/j.jalz.2012.08.008 PMID:23375564

Boyke, J., Driemeyer, J., Gaser, C., Buchel, C., & May, A. (2008). Training-Induced Brain Structure Changes in the Elderly. *The Journal of Neuroscience*, *28*(28), 7031–7035. doi:10.1523/JNEUROSCI.0742-08.2008 PMID:18614670

Bradford, A., Kunik, M. E., Schulz, P., Williams, S. P., & Singh, H. (2009). Missed and Delayed Diagnosis of Dementia in Primary Care. *Alzheimer Disease and Associated Disorders*, *23*(4), 306–314. doi:10.1097/WAD.0b013e3181a6bebc PMID:19568149

Brewer, M. (2014). *Handbook of Research Methods in Social and Personality Psychology*. Retrieved from http://books.google.com/books?id=NRWTAgAAQBAJ&pgis=1

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews. Neuroscience*, *7*(1), 30–40. doi:10.1038/nrn1809 PMID:16371948

Busse, A. L., Gil, G., Santarém, J. M., & Filho, W. J. (2009). Physical activity and cognition in the elderly: A review. *Dementia Neuropsychologia, 3*(3), 204–208. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2010-10986-005&site=ehost-live

Byun, S., & Park, C. (2011). *Serious Game for Cognitive Testing of Elderly HCI International 2011 – Posters' Extended Abstracts* (C. Stephanidis, Ed.). Berlin: Springer Berlin Heidelberg. doi:10.1007/978-3-642-22098-2

Camicioli, R., Howieson, D., Oken, B., Sexton, G., & Kaye, J. (1998). Motor slowing precedes cognitive impairment in the oldest old. *Neurology*, *50*(5), 1496–1498. doi:10.1212/WNL.50.5.1496 PMID:9596020

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Canu, E., McLaren, D. G., Fitzgerald, M. E., Bendlin, B. B., Zoccatelli, G., Alessandrini, F., ... Frisoni, G. B. (2010). Microstructural diffusion changes are independent of macrostructural volume loss in moderate to severe Alzheimer's disease. *Journal of Alzheimer's Disease*, *19*(3), 963–976. doi:10.3233/JAD-2010-1295 PMID:20157252

Cassady, J. C., & Johnson, R. E. (2002). Cognitive Test Anxiety and Academic Performance. *Contemporary Educational Psychology*, *27*(2), 270–295. doi:10.1006/ceps.2001.1094

Chiriac, S., Röll, N., Parada, J., & Rosales, B. (2012). Towards combining validation concepts for short and long-term ambient health monitoring. *2012 6th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2012*. http://doi.org/10.4108/icst.pervasivehealth.2012.248676

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A Meta-Analytic study. *Psychological Science*, *14*(2), 125–130. doi:10.1111/1467-9280.t01-1-01430 PMID:12661673

Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ Pilot Study. *Journal of the American Geriatrics Society*, *55*(10), 1630–1635. doi:10.1111/j.1532-5415.2007.01379.x PMID:17697099

Czaplicka, A., Holyst, J. A., & Sloot, P. M. A. (2013). Noise enhances information transfer in hierarchical networks. *Scientific Reports*, *3*, 1223. doi:10.1038rep01223 PMID:23390574

Delbeuck, X., Collette, F., & Van der Linden, M. (2007). Is Alzheimer's disease a disconnection syndrome? Evidence from a crossmodal audio-visual illusory experiment. *Neuropsychologia*, *45*(14), 3315–3323. doi:10.1016/j.neuropsychologia.2007.05.001 PMID:17765932

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., ... Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *Lancet Neurology*, *6*(8), 734–746. doi:10.1016/S1474-4422(07)70178-3 PMID:17616482

dynanets.org. (2015). Retrieved July 25, 2015, from http://dynanets.org/

Fellgiebel, A., Dellani, P. R., Greverus, D., Scheurich, A., Stoeter, P., & Müller, M. J. (2006). Predicting conversion to dementia in mild cognitive impairment by volumetric and diffusivity measurements of the hippocampus. *Psychiatry Research: Neuroimaging*, *146*(3), 283–287. doi:10.1016/j.pscychresns.2006.01.006 PMID:16530394

Fernández, A., Turrero, A., Zuluaga, P., Gil, P., Maestú, F., Campo, P., & Ortiz, T. (2006). Magnetoencephalographic parietal delta dipole density in mild cognitive impairment: Preliminary results of a method to estimate the risk of developing Alzheimer disease. *Archives of Neurology*, *63*(3), 427–430. doi:10.1001/archneur.63.3.427 PMID:16533970

Flicker, L. (1999). Book: Dementia Reconsidered: the Person Comes First. *BMJ (Clinical Research Ed.)*, *318*(7187), 880–880. doi:10.1136/bmj.318.7187.880a PMID:10092286

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fox, N. C., & Freeborough, P. A. (1997). Brain atrophy progression measured from registered serial MRI: Validation and application to alzheimer's disease. *Journal of Magnetic Resonance Imaging*, *7*(6), 1069–1075. doi:10.1002/jmri.1880070620 PMID:9400851

Frantzidis, C. A., Vivas, A. B., Tsolaki, A., Klados, M. A., Tsolaki, M., & Bamidis, P. D. (2014). Functional disorganization of small-world brain networks in mild Alzheimer's Disease and amnestic Mild Cognitive Impairment: An EEG study using Relative Wavelet Entropy (RWE). *Frontiers in Aging Neuroscience*, *6*, 224. doi:10.3389/fnagi.2014.00224 PMID:25206333

Fratiglioni, L., Wang, H. X., Ericsson, K., Maytan, M., & Winblad, B. (2000). Influence of social network on occurrence of dementia: A community-based longitudinal study. *Lancet*, *355*(9212), 1315–1319. doi:10.1016/S0140-6736(00)02113-9 PMID:10776744

Garcés, P., Angel Pineda-Pardo, J., Canuet, L., Aurtenetxe, S., López, M. E., Marcos, A., ... Maestú, F. (2014). The Default Mode Network is functionally and structurally disrupted in amnestic mild cognitive impairment - a bimodal MEG-DTI study. *NeuroImage. Clinical*, *6*, 214–221. doi:10.1016/j.nicl.2014.09.004 PMID:25379433

Golle, P., & Partridge, K. (2009). Pervasive Computing (Vol. 5538). Berlin: Springer Berlin Heidelberg. http://doi.org/ doi:10.1007/978-3-642-01516-8_26

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(13), 4637–4642. doi:10.1073/pnas.0308627101 PMID:15070770

Hagler, S., Austin, D., Hayes, T. L., Kaye, J., & Pavel, M. (2010). Unobtrusive and ubiquitous in-home monitoring: A methodology for continuous assessment of gait velocity in elders. *IEEE Transactions on Bio-Medical Engineering*, *57*(4), 813–820. doi:10.1109/TBME.2009.2036732 PMID:19932989

Hagler, S., Jimison, H. B., & Pavel, M. (2014). Assessing executive function using a computer game: Computational modeling of cognitive processes. *IEEE Journal of Biomedical and Health Informatics*, *18*(4), 1442–1452. doi:10.1109/JBHI.2014.2299793 PMID:25014944

Hayes, T. L., Abendroth, F., Adami, A., Pavel, M., Zitzelberger, T. A., & Kaye, J. A. (2008). Unobtrusive assessment of activity patterns associated with mild cognitive impairment. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, *4*(6), 395–405. doi:10.1016/j.jalz.2008.07.004 PMID:19012864

Hayes, T. L., Riley, T., Mattek, N., Pavel, M., & Kaye, J. A. (2014). Sleep habits in mild cognitive impairment. *Alzheimer Disease and Associated Disorders*, *28*(2), 145–150. doi:10.1097/WAD.0000000000000010 PMID:24145694

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews. Neuroscience*, *5*(2), 87–96. doi:10.1038/nrn1323 PMID:14735112

Hughes, J. C., Robinson, L., & Volicer, L. (2005). Specialist palliative care in dementia. *BMJ (Clinical Research Ed.)*, *330*(7482), 57–58. doi:10.1136/bmj.330.7482.57 PMID:15637354

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *115*(7), 1490–1505. doi:10.1016/j.clinph.2004.01.001 PMID:15203050

Kantarci, K., Senjem, M. L., Avula, R., Zhang, B., Samikoglu, A. R., Weigand, S. D., ... Jack, C. R. (2011). Diffusion tensor imaging and cognitive function in older adults with no dementia. *Neurology*, *77*(1), 26–34. doi:10.1212/WNL.0b013e31822313dc PMID:21593440

Katzouris, N., Artikis, A., & Paliouras, G. (2014). Event Recognition for Unobtrusive Assisted Living. In A. Likas, K. Blekas, & D. Kalles (Eds.), Lecture Notes in Computer Science: Vol. 8445. *Artificial Intelligence: Methods and Applications* (pp. 237–250). Cham: Springer International Publishing; doi:10.1007/978-3-319-07064-3

Konstantinidis, E., Antoniou, P., Billis, A., Bamparopoulos, G., Costas, P., & Bamidis, P. (2014). *Universal Access in Human-Computer Interaction. Aging and Assistive Environments* (C. Stephanidis & M. Antona, Eds.). Cham: Springer International Publishing. doi:10.1007/978-3-319-07446-7

Konstantinidis, E., Billis, A., Mouzakidis, C., Zilidou, V., Antoniou, P., & Bamidis, P. (2014). Design, implementation and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens. *IEEE Journal of Biomedical and Health Informatics*, 1–1. http://doi.org/ doi:10.1109/JBHI.2014.2378814

Konstantinidis, E. I., Antoniou, P. E., Bamparopoulos, G., & Bamidis, P. D. (2014). A lightweight framework for transparent cross platform communication of controller data in ambient assisted living environments. *Information Sciences*, *300*, 124–139. doi:10.1016/j.ins.2014.10.070

Konstantinidis, E. I., Bamidis, P. D., Billis, A. S., Kartsidis, P., & Papageorgiou, S. G. (2015). Physical training in-game metrics for cognitive assessment: Evidence from extended trials with the FitForAll exergame platform. *PLoS ONE*.

Kotagal, V., Langa, K. M., Plassman, B. L., Fisher, G. G., Giordani, B. J., Wallace, R. B., ... Foster, N. L. (2015). Factors associated with cognitive evaluations in the United States. *Neurology*, *84*(1), 64–71. doi:10.1212/WNL.0000000000001096 PMID:25428689

Leung, P., Orrell, M., & Orgeta, V. (2015). Social support group interventions in people with dementia and mild cognitive impairment: A systematic review of the literature. *International Journal of Geriatric Psychiatry*, *30*(1), 1–9. doi:10.1002/gps.4166 PMID:24990344

Liu, S., Shen, Z., Mei, J., & Ji, J. (2013). Parkinson's Disease Predictive Analytics through a Pad Game Based on Personal Data. *Int. J. Inf. Technol, 19*(2).

Liu, Y., Wang, K., Yu, C., He, Y., Zhou, Y., Liang, M., ... Jiang, T. (2008). Regional homogeneity, functional connectivity and imaging markers of Alzheimer's disease: A review of resting-state fMRI studies. *Neuropsychologia*, *46*(6), 1648–1656. doi:10.1016/j.neuropsychologia.2008.01.027 PMID:18346763

Lyketsos, C. G., Lopez, O., Jones, B., Fitzpatrick, A. L., Breitner, J., & DeKosky, S. (2002). Prevalence of Neuropsychiatric Symptoms in Dementia and Mild Cognitive Impairment. *Journal of the American Medical Association*, *288*(12), 1475. doi:10.1001/jama.288.12.1475 PMID:12243634

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Lyons, B. E., Austin, D., Seelye, A., Petersen, J., Yeargers, J., & Riley, T., … Kaye, J. a. (2015). Pervasive Computing Technologies to Continuously Assess Alzheimerâ€™s Disease Progression and Intervention Efficacy. *Frontiers in Aging Neuroscience*, *7*(June), 1–14. doi:10.3389/fnagi.2015.00102 PMID:25653617

Mahncke, H. W., Connor, B. B., Appelman, J., Ahsanuddin, O. N., Hardy, J. L., Wood, R. A., … Merzenich, M. M. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: a randomized, controlled study. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(33), 12523–8. http://doi.org/10.1073/pnas.0605194103

McCormick, W. C., Kukull, W. A., van Belle, G., Bowen, J. D., Teri, L., & Larson, E. B. (1994). Symptom Patterns and Comorbidity in the Early Stages of Alzheimer's Disease. *Journal of the American Geriatrics Society*, *42*(5), 517–521. doi:10.1111/j.1532-5415.1994.tb04974.x PMID:8176147

McKanna, J. A., Jimison, H., & Pavel, M. (2009). Divided attention in computer game play: analysis utilizing unobtrusive health monitoring. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2009*. http://doi.org/10.1109/IEMBS.2009.5334662

McKeith, I. G., Dickson, D. W., Lowe, J., Emre, M., O'Brien, J. T., Feldman, H., ... Yamada, M. (2005). Diagnosis and management of dementia with Lewy bodies: Third report of the DLB Consortium. *Neurology*, *65*(12), 1863–1872. doi:10.1212/01.wnl.0000187889.17253.b1 PMID:16237129

Mehta, K. M., Yaffe, K., & Covinsky, K. E. (2002). Cognitive Impairment, Depressive Symptoms, and Functional Decline in Older People. *Journal of the American Geriatrics Society*, *50*(6), 1045–1050. doi:10.1046/j.1532-5415.2002.50259.x PMID:12110064

Mevel, K., Chételat, G., Eustache, F., & Desgranges, B. (2011). The default mode network in healthy aging and Alzheimer's disease. *International Journal of Alzheimer's Disease*, *535816*. doi:10.4061/2011/535816 PMID:21760988

Mihailidis, A. (2008). *Technology and Aging: Selected Papers from the 2007 International Conference on Technology and Aging*. Retrieved from https://books.google.com/books?hl=el&lr=&id=OSCScvzDPrwC&pgis=1

Monekosso, D. N., & Remagnino, P. (2010). Behavior Analysis for Assisted Living. *IEEE Transactions on Automation Science and Engineering*, *7*(4), 879–886. doi:10.1109/TASE.2010.2049840

Oishi, K., Mielke, M. M., Albert, M., Lyketsos, C. G., & Mori, S. (2012). The fornix sign: A potential sign for Alzheimer's disease based on diffusion tensor imaging. *Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging*, *22*(4), 365–374. doi:10.1111/j.1552-6569.2011.00633.x PMID:21848679

Panagiotis, D., Bamidis, Tarnanas, I., Tsolaki, M., & Hadjileontiadis, L. (2015). *Handbook of Research on Innovations in Database Technologies and Applications: Current and Future Trends*. Retrieved July 29, 2015, from http://www.igi-global.com/book/handbook-research-innovations-diagnosis-treatment/121164

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Papageorgiou, E. I., Billis, A. S., Frantzidis, C., Konstantinidis, E. I., & Bamidis, P. D. (2013). A preliminary fuzzy cognitive map - based desicion support tool for geriatric depression assessment. In *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1–8). IEEE. http://doi.org/10.1109/FUZZ-IEEE.2013.6622485

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2014). Multisensory integration during short-term music reading training enhances both uni- and multisensory cortical processing. *Journal of Cognitive Neuroscience*, *26*(10), 2224–2238. doi:10.1162/jocn_a_00620 PMID:24669793

Quax, R., Apolloni, A., & Sloot, P. M. A. (2013). Towards understanding the behavior of physical systems using information theory. *The European Physical Journal. Special Topics*, *222*(6), 1389–1401. doi:10.1140/epjst/e2013-01933-9

Quax, R., Kandhai, D., & Sloot, P. M. A. (2013). Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series. *Scientific Reports*, *3*, 1898. doi:10.1038rep01898 PMID:23719567

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–682. doi:10.1073/pnas.98.2.676 PMID:11209064

Rashidi, P., & Mihailidis, A. (2013). A Survey on Ambient-Assisted Living Tools for Older Adults. *IEEE Journal of Biomedical and Health Informatics*, *17*(3), 579–590. doi:10.1109/JBHI.2012.2234129 PMID:24592460

Reiman, E. M., Quiroz, Y. T., Fleisher, A. S., Chen, K., Velez-Pardo, C., Jimenez-Del-Rio, M., ... Lopera, F. (2012). Brain imaging and fluid biomarker analysis in young adults at genetic risk for autosomal dominant Alzheimer's disease in the presenilin 1 E280A kindred: A case-control study. *Lancet Neurology*, *11*(12), 1048–1056. doi:10.1016/S1474-4422(12)70228-4 PMID:23137948

Risacher, S. L., Saykin, A. J., West, J. D., Shen, L., Firpi, H. A., & McDonald, B. C. (2009). Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Current Alzheimer Research*, *6*(4), 347–361. doi:10.2174/156720509788929273 PMID:19689234

Robert, P. H., König, A., Amieva, H., Andrieu, S., Bremond, F., Bullock, R., ... Manera, V. (2014). Recommendations for the use of Serious Games in people with Alzheimer's Disease, related disorders and frailty. *Frontiers in Aging Neuroscience*, *6*, 54. doi:10.3389/fnagi.2014.00054 PMID:24715864

Röcker, C., Hinske, S., & Magerkurth, C. (2007). Universal Access in Human-Computer Interaction. Ambient Interaction (Vol. 4555). Berlin: Springer Berlin Heidelberg. http://doi.org/ doi:10.1007/978-3-540-73281-5

Rogalski, E., Stebbins, G. T., Barnes, C. A., Murphy, C. M., Stoub, T. R., George, S., ... deToledo-Morrell, L. (2012). Age-related changes in parahippocampal white matter integrity: A diffusion tensor imaging study. *Neuropsychologia*, *50*(8), 1759–1765. doi:10.1016/j.neuropsychologia.2012.03.033 PMID:22561887

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Scarmeas, N., Levy, G., Tang, M.-X., Manly, J., & Stern, Y. (2001). Influence of leisure activity on the incidence of Alzheimer's Disease. *Neurology*, *57*(12), 2236–2242. doi:10.1212/WNL.57.12.2236 PMID:11756603

Seelye, A., Mattek, N., Howieson, D., Riley, T., Wild, K., & Kaye, J. (2015). The Impact of Sleep on Neuropsychological Performance in Cognitively Intact Older Adults Using a Novel In-Home Sensor-Based Sleep Assessment Approach. *The Clinical Neuropsychologist*, *29*(1), 53–66. doi:10.1080/13854046.2015.1005139 PMID:25642948

Sheline, Y. I., Morris, J. C., Snyder, A. Z., Price, J. L., Yan, Z., D'Angelo, G., ... Mintun, M. A. (2010). APOE4 allele disrupts resting state fMRI connectivity in the absence of amyloid plaques or decreased CSF Aβ42. *The Journal of Neuroscience*, *30*(50), 17035–17040. doi:10.1523/JNEUROSCI.3987-10.2010 PMID:21159973

Shi, F., Liu, B., Zhou, Y., Yu, C., & Jiang, T. (2009). Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: Meta-analyses of MRI studies. *Hippocampus*, *19*(11), 1055–1064. doi:10.1002/hipo.20573 PMID:19309039

Staiano, A. E., & Calvert, S. L. (2011). The promise of exergames as tools to measure physical health. *Entertainment Computing*, *2*(1), 17–21. doi:10.1016/j.entcom.2011.03.008 PMID:23378860

Storf, H., Becker, M., & Riedl, M. (2009). Rule-based activity recognition framework: Challenges, technique and learning. In *Proceedings of the 3d International ICST Conference on Pervasive Computing Technologies for Healthcare* (pp. 1–7). ICST. http://doi.org/10.4108/ICST.PERVASIVEHEALTH2009.6108

Styliadis, C., Kartsidis, P., & Paraskevopoulos, E. (n.d.). *Neuroimaging Approaches for Elderly Studies*. IGI Global. Retrieved from http://www.igi-global.com/chapter/neuroimaging-approaches-for-elderly-studies/129268

Styliadis, C., Kartsidis, P., Paraskevopoulos, E., Ioannides, A. A., & Bamidis, P. D. (2015). Neuroplastic Effects of Combined Computerized Physical and Cognitive Training in Elderly Individuals at Risk for Dementia: An eLORETA Controlled Study on Resting States. *Neural Plasticity*, *12*(2). Retrieved from http://doi.org/doi:10.1155/2015/172192 PMID:25945260

Tardif, S., & Simard, M. (2011). Cognitive stimulation programs in healthy elderly: A review. *International Journal of Alzheimer's Disease*, *378934*. doi:10.4061/2011/378934 PMID:21876829

Tarnanas, I., Schlee, W., Tsolaki, M., Müri, R., Mosimann, U., & Nef, T. (2013). Ecological validity of virtual reality daily living activities screening for early dementia: Longitudinal study. *JMIR Serious Games*, *1*(1), e1. doi:10.2196/games.2778 PMID:25658491

Tractenberg, R. E., Singer, C. M., & Kaye, J. (2005). Symptoms of sleep disturbance in persons with Alzheimer's disease and normal elderly. *Journal of Sleep Research*, *14*(2), 177–185. doi:10.1111/j.1365-2869.2005.00445.x PMID:15910515

Wang, L., Zang, Y., He, Y., Liang, M., Zhang, X., Tian, L., ... Li, K. (2006). Changes in hippocampal connectivity in the early stages of Alzheimer's disease: Evidence from resting state fMRI. *NeuroImage*, *31*(2), 496–504. doi:10.1016/j.neuroimage.2005.12.033 PMID:16473024

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Wang, X. (2014). Subjective well-being associated with size of social network and social support of elderly. *Journal of Health Psychology*. http://doi.org/ doi:10.1177/1359105314544136

WHO. (2012). *Dementia: a public health priority*. Retrieved from http://www.who.int/mental_health/publications/dementia_report_2012/en/

Xefteris, S., Andronikou, V., Tserpes, K., & Varvarigou, T. (2010). Case-based approach using behavioural biometrics aimed at Assisted Living. *Journal of Ambient Intelligence and Humanized Computing*, *2*(2), 73–80. doi:10.100712652-010-0029-8

Xefteris, S., Baboshin, A., Tserpes, K., Androulidakis, A., Glickman, Y., Varvarigou, T., … D'Andria, F. (2011). Enabling risk assessment and analysis by event detection in dementia patients using a reconfigurable rule set. In *Proceedings of the 4th International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '11* (p. 1). New York: ACM Press. http://doi.org/10.1145/2141622.2141678

Yang, G., Xie, L., Mantysalo, M., Zhou, X., Pang, Z., Xu, L., … Zheng, L.-R. (2014). A Health-IoT Platform Based on the Integration of Intelligent Packaging, Unobtrusive Bio-Sensor, and Intelligent Medicine Box. *IEEE Transactions on Industrial Informatics, 10*(4), 2180–2191. http://doi.org/ doi:10.1109/TII.2014.2307795

Yang, S. (2010). Defining Exergames & Exergaming. *Proceedings of Meaningful Play*, *2010*, 1–17. Retrieved from http://meaningfulplay.msu.edu/proceedings2010/mp2010_paper_63.pdf

Zarrabi, N., Prosperi, M. C. F., Belleman, R. G., Di Giambenedetto, S., Fabbiani, M., De Luca, A., & Sloot, P. M. A. (2013). Combining social and genetic networks to study HIV transmission in mixing risk groups. *The European Physical Journal. Special Topics*, *222*(6), 1377–1387. doi:10.1140/epjst/e2013-01932-x

This research was previously published in Improving the Quality of Life for Dementia Patients through Progressive Detection, Treatment, and Care edited by Jinglong Wu; pages 74-106; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 99
Opportunities and Challenges of Big Data in Healthcare

Wafaa Faisal Mukhtar
Sudan University of Science and Technology, Sudan

Eltayeb Salih Abuelyaman
University of Dammam, Saudi Arabia

ABSTRACT

Healthcare big data streams from multiple information sources at an alarming volume, velocity, and variety. The challenge that faces the healthcare industry is extracting meaningful value from such sources. This chapter investigates the diversity and forms of data in the healthcare sector, reviews the methods used to search and analyze these data throughout the past years, and the use of machine learning and data mining techniques to mine useful knowledge from such data. The chapter will also highlight innovations of particular systems and tools which spot the fine approaches for different healthcare data, raise the standard of care and recap the tools and data collection methods. The authors emphasize some of ethical issues regarding processing these records and some data privacy issues.

INTRODUCTION

We live in the age of Big Data, where every single data entry to any computer at home or office is being recorded in data centers and processed later in order to get valuable information or knowledge. IDC (International Data Corporation) defines Big Data technologies as a new generation of technologies and architectures, designed to extract economical *value* from very large *volumes* of a wide *variety* of data produced every day, by enabling high *velocity* capture, discovery, and/or analysis (Gantz & Reinsel, 2011). Other institutes and companies such as SAS, IBM and Oracle have other dimensions for Big Data, but they all address same concepts.

DOI: 10.4018/978-1-7998-1204-3.ch099

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Several organizations create very large databases, such as consumer data and transaction histories of sales records, patient records, images at hospitals, fingerprints and DNA samples at crime scenes, and data acquired via satellites or surveillance cameras. It is apparent that there is a huge gap between the amount of available information and its corresponding knowledge, which gives an opportunity for extracting hidden knowledge.

Healthcare can be considered a major sector for producing big data. Healthcare data have many dimensions such as personal information regarding dates of birth, nationality, sex, location and other information about income and education plus other specific data about history of the patient with medical centers and/or disease. It is generated by everything around us at all times in intensive care units and in surgical operation room, inward_from every digital process in systems or sensors.

The aim of this chapter is to discuss the notion of big data when applied to the healthcare sector. The next section reveals how and why big data are represented in health care. Big Data's Vs will be emphasized for healthcare data in the third section. The fourth section highlights different techniques that are well suited for intelligent data analysis. Section five illustrates the challenges facing Big Data and the opportunities available for valuing analysis, and some ethical, legal and social concerns about manipulating healthcare data specifically. Some implications are contending in the future directions showing how they can play an important role in the overall improvement of the healthcare industry.

BACKGROUND

Medical data are at once the most rewarding and challenging of all biological data. For decades everyone was infatuated by the liability of keeping every record and collecting any possible information about everything in their life. The healthcare industry has also experienced these practices about generating and keeping large amounts of data driven by record keeping at physicians' clinics, which is referred to as patient records. This includes forms filled by the patient regarding his/her personal information and oral examination recorded by physicians during visits. Other forms of checkups, different laboratory examinations, and CT scan as well as X-ray images are also kept in hospital's emergency room when examining patients. Moreover, data about compliance & regulatory requirements, and patient care is also evolving from national and international organizations that monitor and administer the healthcare industry.

Electronic health records have experienced several studies. Drug safety study (Trifirò et al., 2009) investigated adverse drug reactions with other diseases, in (Jensen, Jensen, & Brunak, 2012), they combined the HER with the genetic data to reveal gene-disease association, (Almodaifer, Hafez, & Mathkour, 2011) discovered the interesting and concise medical rules for prediction purpose to assist the medical decision makers.

Medical diagnosis researches have proven a great success, because the data about the disease and the patient under examination is always available. In fact the medical diagnostic knowledge can be automatically derived from the description of cases solved in the past. (Kumar, Sathyadevi, & Sivanesh, 2011) proposed using an intelligent clinical decision support system to assist physicians in diagnosing. An automatic diagnosis system was presented in (Karabatak & Ince, 2009b). Soni & Ansari, 2011; Kharya, 2012; Huang, Chen, & Lee, 2007; (Ha, 2011) (Kononenko, 2001) summarized several machine learning techniques used for classifying diseases such as naïve Bayesian and neural networks, his work also highlighted the specific requirements for good performing machine learning algorithms in solving medical diagnostic tasks.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Image parameterization observed in medical records is very useful in quality control, identification, image grouping, surveillance, image storage and retrieval, and image querying. Texture classification is closely related to diagnostic process to diagnose coronary artery disease (Kukar, Kononenko, & Grošelj, 2011).

Big Data V's

The V's that characterizes Big Data (volume, variety, velocity, and value) will be identified with exceptional respect to healthcare sector.

Volume

The volume of generated and stored data is extremely expanding every year, and there is an evolution amongst the country level, in digitizing all transactions. Electronic health records (EHR) are stored in databases since the eighties of the past century. These records are generated when a patient gets medication service at a physician clinic, hospital or even from the pharmacy. In the past there was no collaboration between these different entities which resulted in duplication and inconsistence of patient records. According to data.worldbank.org the world population is about 7 billion, where about 15% lived in developed countries. EHRs types, which vary from any, basic to fully functional, were adopted in these developing countries, and there are plans to establish health systems in other urban areas, which describe the volume of records in anticipation of processing. Moreover, data also is generated during laboratory examinations of blood and other fluids, images of different types and are all kept recently in electronic formats. These records are kept not only for living people, but also for the people who died several years ago. Data about drug specifications, formulations, interaction, side effects and safety are all kept as pharmaceutical documentary records. Billing, financial transactions and insurance claims are also archived over years.

Variety

The variability and complexity of healthcare data come from different resources that provide the data in different formats. Patient records by themselves can have different forms, since some clinics or hospitals can adopt a similar format, although some private clinics may not. Some records may contain images with different size and dimensions like X-Ray, CT Scan or MRI which increase the complexity of handling these data. Laboratory tests may also differ between clinics and hospitals. The prescription advised by the physician can be kept as text or scanned image. In the operation or intensive care units, nurses and doctors may keep observatory text notes and examination checks.

Survey information on population health, such as causes of deaths, community health, and health care systems are collected in Population health metrics. National health organizations control enrollment into health insurance plans in order to administer health care delivery. Huge statistics on public health databases and registries concerning births, deaths, disease incidence, health event case reports, demographics and more are available at United Nations organizations.(HealthData.gov, 2014)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Velocity

Medical databases are constantly updated in critical health conditions. Dealing with data from intensive care units (ICUs) and operating rooms (ORs) are a challenge in hospitals. Continuous and high priority monitoring add new records and tests in every second. There is a necessity to update the databases since doctors and nurses write progress notes. Drug prescription and delivery is changed and charted several times a day. Monitors generate large amounts of data on a continuous basis, as well as therapeutic devices such as mechanical ventilators, syringe- and infusion pumps for drug and fluid administration, or renal replacement therapy machines. Blood samples for ICU occupants are drawn for laboratory analysis several times a day, and microbiology sampling occurs several times a week (Meyfroidt, Güiza, Ramon, & Bruynooghe, 2009). At Home monitoring systems for critical disease patients through sensing devices such as heart monitors, ventilators, cell phones accelerometer or wearable devices provide real-time data.

Value

It is, however, the value, veracity, or 'data assurance' that should generate an improvement in the healthcare industry. In order to benefit from the mass of collected data; researchers, practitioners and policy makers should analyze these data and get implication which will strengthen decision making and improve healthcare management. Organizations should be able to analyze health problems through investigating health claims, get a more complete view of health processes and their interactions by analyzing patients' records, evaluating a greater number of scenarios quicker and more cheaply to best achieve operational efficiency at ICU units and operation rooms, employees satisfaction, and risk reduction for healthcare organizations across the nation. Drug safety researchers need to get new experiences into medication drug collaborations that have not been completely comprehended or anticipated during clinical trials, this can be achieved by association analysis at the pharmacology level and public health researchers. (Silvia & Claps, 2013)

The healthcare industry is moving from reporting truths to disclosure of experiences, getting to be information driven. Converting data and information into actionable insights by several procedures of data analytics to prevent inefficiencies and adapt workflows for improved healthcare outcomes across the end-to-end patient journey will be paramount. It will empower better joint effort among clinical experts, engaging general health authorities to supervise the individualized gathering anticipation movement, and providers to implement new industry models, pushing performance by reward could set scales for financing healthcare services, making it available extensively for people. (Silvia & Claps, 2013)

Eventually, it is the value that will make Big Data a genuine power of progress in healthcare, ensuring value while preserving quality.

HEALTHCARE DATA ANALYSIS

In addition to the high volume and velocity, the complexity and variety of big data, Gartner IT dictionary has added that big data also demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Gandomi & Haider, 2015). Data can be described as big if the growth of data far exceeds the efforts done to understand it.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

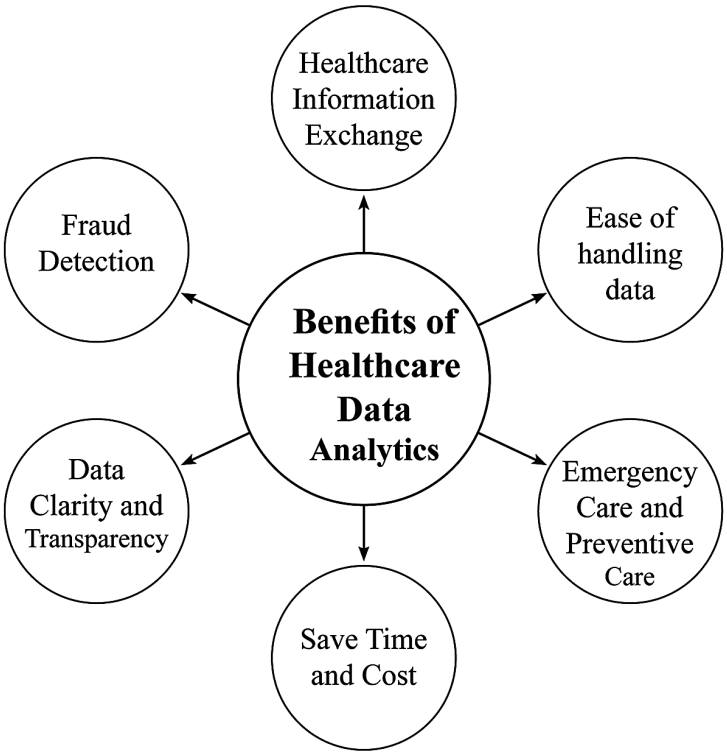
Healthcare data analysis can be advantageous in building predictive modeling that can lower attrition and results in fine procedures in clinical operations and drugs safety. For public health improvement, analysis of largely produced medical data can help in finding disease patterns and tracking epidemic diseases, turn data into actionable evidence to improve services and prevent fraud and crises (Raghupathi & Raghupathi, 2014). Figure 1 illustrates that different employees in the healthcare sector will aggravate better choices concerning strategy making, inventing new instruments to capture data, new strategies for diagnosing and prognosis, developing patient engagement, optimizing operations, putting off fraud as well as threats, and profiting by introducing a new supply of income.

Due to the variety of data, which holds images, forms, databases and surveys, analysis techniques that deals especially with large amount of data is adopted. Machine learning algorithms and data mining techniques proved to be the best practices which aids and supports the process of decision making.

Machine Learning

Statistics for decades was the best analytical tool for medical students and researchers. It was used for survey analysis, generating reports and representing charts, comparative studies about different health care distribution and disease/drug relationship. Due to the fast evolution in computation and computers, machine learning became the new solution. A major subfield of artificial intelligence based on statistical foundation, machine learning is a multidisciplinary field involving information theory, philosophy, neurobiology and other fields.

Figure 1. Benefits of healthcare data analytics
Dergurahian, 2011.



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The fast growth of databases and the lack of knowledge mining, challenged researchers to benefit from the low-priced computation power. Big Data investigation requires enormous programming effort, which can be simplified using intelligent systems that support learning making it simple and comprehensible, a program can be automatically driven. Machine learning algorithms are probably the most valuable candidates for health care data analytics because they can deal with different and large amounts of data integrated from different sources and they also rely explicitly on previous domain knowledge or experience in the analysis, providing several programs for intelligent data analysis.

Supervised learning algorithms discover automatically a knowledge model as a result of the learning process that provides a description of the data explored. Support Vector Machine (SVM), Decision Tree (DT), Artificial Neural Networks (ANNs), Regression, K-nearest neighborhood (KNN), Bayesian Networks (BN) and many more algorithms were considered for machine learning to classify and organize big data easily. The aim of the knowledge model is to predict the value of the target attribute for new unseen patterns.

Another type of learning is unsupervised learning, which analyzes the information or data with unknown target variable, building the model which solely describes the data analyzed. The goal of the process is to build a model that describes interesting regularities in the data.

These learning techniques are mostly used in medical diagnosis, and classifying patients accordingly to different attributes available in their medical records (Hariz, Adnan, Husain, Aini, & Rashid, 2012; Kononenko, 2001; Meyfroidt et al., 2009; Yang & Wu, 2006), Lots of researches regarding machine learning in medical data and bioinformatics are available in (Larrañaga et al., 2006; Prompramote, Chen, & Chen, 2005; Lavrac, 1999; Lavrač, 1999).

Data Mining

As the gap between produced data its understanding is widening, it is high time to look for potentially useful information which lies hidden in all these hypes of data or information. Data mining or Knowledge Discovery from Databases (KDD) helps extracting such patterns (Witten, Frank, & Hall, 2011). Data mining has been used earlier in market analysis, financial data analysis, business management, space exploration and proved to be the best solution for various domains especially medical data analysis (Tomar & Agarwal, 2013).

Healthcare industry can benefit from using data mining techniques to take considerable actions regarding: (Tomar & Agarwal, 2013)

1. Effective management of Hospital resource,
2. Hospital Ranking,
3. Better Customer Relation,
4. Hospital Infection Contro,
5. Smarter Treatment Techniques,
6. Improved Patient care,
7. Decrease Insurance Fraud,
8. Recognize High-Risk Patients,
9. Health Policy Planning.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The nature, complexity and growth of the data in healthcare sector, brings medical data mining to the edge of new technologies, proposing a powerful tool for analyzing and extracting important knowledge.

The knowledge can help building models that might support prediction of diseases or recommend drug, understanding the nature or describe associations of the data been collected. Applying data mining concepts, such as prediction, description, summarization and visualization to the healthcare big data improved the industry, saving time, money and life (Jacob, 2012).

Medical data may take tabular form or standard format *m.n* array where *m* represents the size of the data in rows and *n* represents the features or attributes in which data is described. Classification techniques, built across supervised learning techniques, dominates other data mining techniques because it is can assist in the future prediction by building a model of a classifier. Practitioners may seek for quantitative or qualitative description of the data, grouping patients with similar features or having similar disease type, using association analysis and clustering techniques.

Natural language processing and text mining can be applied for medical records or drug prescription, applying different techniques for classification, clustering or even discover new associations. Images can undergo digital image processing and then apply image mining to extract useful and hidden patterns.

Medical diagnosis systems can be developed using mining techniques, investigate symptoms and patient medical history to detect certain disease, classify image as malignant or benign for tumer detection, find explanation for causes of diseases through descriptive analysis such as association analysis, clustering and sequential pattern mining. Data mining patterns could be analyzed to prevent developing of diabetes or high blood pressure for example.

Challenges and Solutions

The significance and value of Big data lies in the storage facilities, functioning processes and procedures used for sharing and capturing data and tools used to analyze and search very large data sets.

The first challenge that faces medical data mining is the lack of a standardized clinical vocabulary (Jacob, 2012). The mass of healthcare data is buried in heterogeneous databases scattered throughout many institution, tied in with a lot of missing, distorted, conflicting, and non-homogenous data in different formats in diverse data sources, thus most probably might be ignored during analysis. The stored information becomes less useful if it is not available in an easily comprehensible format. Nevertheless, large portion of this gathered data will be unprocessed, for instance, video streaming data are deleted after a week, instead some sort of archiving should be adopted every period and some arrangement of knowledge extraction, such as using it for performance auditing (Raghupathi & Raghupathi, 2014).

Bio/medical big data has major challenges, computation power and data analysis. Computation cost while analyzing the biological data especially after using the next generation sequencing technology that produces large size of data. Bio/medical cloud services have developed integrated tools and biological databases to formulate complex systems (Chae et al., 2013).

Data Mining Tools

Powerful automated tools to analye, visualize and interpret useful information from different data sources will be benificial. Many manufacturers support such tools to discover patterns involving methods using the artificial intelligence, machine learning and statistical analysis. KD nuggets offered many Analytics and Data Mining Suites, supporting classification, clustering, data preparation, visualization, and other tasks,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

it also provides a list of specific healthcare analytics solutions (Kdnuggets, 2015). Microsoft developed the SQL Server Analysis Services platform to build high performance multidimensional and tabular analytical models for interactive data analysis, reporting, and visualization (Russo, Ferrari, & Webb, 2012). DB2® Intelligent Miner(TM), was developed at IBM to provide Intelligent Miner analytics such as modeling, scoring and visualization (Ibm, 2014). DB2® was integrated with Clementine data mining workbench since 2004 (Ibm, 2007). SAS and Intel provide an analytical solution as well, SAS® Enterprise Miner™, that can handle accurately large amounts of data and produce predictive and descriptive models (Sas, 2014). Oracle Data Miner deploy data mining models inside the relational database environment (Tamayo, 2005). There are also open source tools for data mining, and the six powerful and mostly used are RapidMiner, Weka, R, Orange, Knime and Natural Language Toolkit(NLTK) (Goopta, 2013).

Novel Frameworks

Practitioners need systems and tools in order to capture and store this information, researchers need optimal processing power, analytics capabilities and skills to search and share the information. Big Data organizations start investing in the development of new infrastructures for running operational, real-time applications that can cope with the enormous and various amounts of information produced.

Yahoo after working with Nutch and Google released Hadoop as an open-source project for reliable, scalable and distributed computing in 2008, but it is now supported by Apache Software Foundation (ASF), it can perform two tasks: data storage and processing (Sas.com, 2014). The Hadoop platform comprise of three main components;

- **Hadoop Distributed File System (HDFS):** A Java based distributed file system that provides high-throughput access to application data that can store all kinds of data without prior organization.
- **Hadoop YARN:** A resource management framework for job scheduling, handling resource requests from distributed applications and cluster resource management.
- **Hadoop MapReduce:** A software programming model for parallel processing of large data sets. It is also accompanied by some other products for handling scalable data bases, warehousing, machine learning, data mining and many more(Apache.org, 2015).

Traditional relational databases can no longer handle and get value from healthcare big data, JustOneDB is a new high performance NoSql relational databases suitable for Big Data storage, load, and fast retrieval of queries and predictive analytics for healthcare industry (JustOneDB, 2015).

Recently bio/medical cloud services have begun evolving from simple and primitive applications to offering integrated and complex systems. The bioinformatics cloud service provides systems such as Galaxy, DNAnexus, Crossbow, SoapSNP and many others, not only without worrying about the manual system installation and data integration processes, but also with reduced data uploading/transferring time. (Chae et al., 2013)

High speed transfer of large volumes of data is achieved byAspera and Intel. Aspera's patented fasp* transport technology is designed and implemented to have no theoretical throughput limit on Intel® Xeon® processor E5-2600 product family in Virtualized Environments (Intel, 2013).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Ethical, Legal, and Social Issues

Medical data contains sensitive, private and confidential information about patients, accordingly the public may fear of data misuse, abuse and disclosure. Data ownership is an open question, since data is collected at different healthcare position, processed mainly by third party. Patient might hesitant against the release of information, although data analysts need to combine various records sometimes from another society. Records are shared and linked more widely nowadays than before while working on the open internet platforms. Security of patient's data during electronical transfer through the internet should be handled within ownership. Furthermore, medical data mining can drive the fear of lawsuits directed against healthcare providers. Patients may raise claims about unnecessary tests performed, confident questions asked and particular conclusion derived unconsciously.

RECOMMENDATIONS AND FUTURE RESEARCH DIRECTIONS

Healthcare Big Data by its different dimensions provide big opportunities for researchers and healthcare practitioners. Everyday activities are moving online, shared and linked to medical transactions or others. Data gathered during any medical transaction should be secured, encrypted or anonymized. Easily accessed and explored data storage facilities will have made it easy to govern and direct plans and strategies to provide the best quality healthcare services. Medical data mining can practice new algorithms for real time data analysis.

Healthcare records have to be standardized to avoid the future data cleansing and inconvenient data. Control over medical data ownership will minimize legal suits against health institutes. Improving care quality consequently will produce value to identify modern approaches for healthcare administration.

CONCLUSION

The chapter went through the advent of Big Data in Healthcare industry and emphasized its four dimensions. The critical manner of processing such data which are of various forms in big volumes and rushing in high velocity in time is clarified. Extracting knowledge or value is a need using new methods and technologies available. Some of the new storage and transfer platforms as well as high performance computing, on the cloud were mentioned. Privacy and confidentiality issues were discussed when trying to merge or analyze such data.

Big data in the health industry is still promising, offering new opportunities for new data storage, transfer and analysis frameworks, and try to get from the exploding information gathered.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Almodaifer, G., Hafez, A., & Mathkour, H. (2011). Discovering medical association rules from medical datasets. *Proceedings of the 2011 IEEE International Symposium on IT in Medicine and Education ITME 2011* (Vol. 2, pp. 43–47). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6132053

Apache.org. (2015). *Welcome to Apache Hadoop.* Retrieved from http://hadoop.apache.org/

Chae, H., Jung, I., Lee, H., Marru, S., Lee, S.-W., & Kim, S. (2013). Bio and health informatics meets cloud : BioVLab as an example. *Health Information Science and Systems*, *1*(1), 6. doi:10.1186/2047-2501-1-6 PMID:25825658

Dergurahian, J. (2011). *Analytics: Moving health care forward. techtarget*. Retrieved from http://search-healthit.techtarget.com/report/Analytics-Moving-health-care-forward

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. doi:10.1016/j.ijinfomgt.2014.10.007

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *White Paper, IDC*, June, 1–12. Retrieved from http://www.itu.dk/people/rkva/2011-Fall-SMA/readings/ExtractingValuefromChaos.pdf

Goopta, C. (2013). *Six of the Best Open Source Data Mining Tools - The New Stack*. Retrieved from http://thenewstack.io/six-of-the-best-open-source-data-mining-tools/

Ha, S. H. S. (2011). Medical Domain Knowledge and Associative Classification Rules in Diagnosis. *International Journal of Knowledge Discovery in Bioinformatics*, *2*(1), 60–73. doi:10.4018/jkdb.2011010104

Hariz, M., Adnan, M., Husain, W., Aini, N., & Rashid, A. (2012). Data Mining for Medical Systems : A Review, (Figure 2), 978–981.

HealthData.gov. (2014). *HealthData.gov*. Retrieved from https://healthdata.gov

Huang, M.-J., Chen, M.-Y., & Lee, S.-C. (2007). Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis. *Expert Systems with Applications*, *32*(3), 856–867. doi:10.1016/j.eswa.2006.01.038

IBM. (2007). Clementine 11. Retrieved from http://www-304.ibm.com/partnerworld/gsd/solutiondetails.do?solution=10387&expand=true

IBM. (2014). DB2 Intelligent Miner. Retrieved from http://www-01.ibm.com/support/knowledgecenter/SSEPGG_9.1.0/com.ibm.im.overview.doc/c_ibm_db2_intelligent_miner_modeling.html

Intel. (2013). Big Data Technologies for Ultra-High-Speed Data Transfer in Life Sciences.

Jacob, S. G. (2012). Data Mining in Clinical Data Sets: A Review. *International Journal of Applied Information Systems*, *4*(6), 15–26. doi:10.5120/ijais12-450774

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: Towards better research applications and clinical care. *Nature Reviews. Genetics*, *13*(6), 395–405. doi:10.1038/nrg3208 PMID:22549152

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

JustOneDB. (2015). *JustOneDB Big Data Analytics*. Retrieved from http://www.justonedb.com/products/justonedb/

Karabatak, M., & Ince, M. C. (2009a). A new feature selection method based on association rules for diagnosis of erythemato-squamous diseases. *Expert Systems with Applications*, *36*(10), 12500–12505. doi:10.1016/j.eswa.2009.04.073

Karabatak, M., & Ince, M. C. (2009b). An expert system for detection of breast cancer based on association rules and neural network. *Expert Systems with Applications*, *36*(2), 3465–3469. doi:10.1016/j.eswa.2008.02.064

Kdnuggets. (2015). Healthcare Analytics Solutions. Retrieved from http://www.kdnuggets.com/solutions/healthcare-analytics.html

Kharya, S. (2012). Using data mining techniques for diagnosis and prognosis of cancer disease. *International Journal of Computer Science, Engineering and Information Technology, 2*(2), 55–66. Retrieved from http://arxiv.org/abs/1205.1923

Kononenko, I. (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine*, *23*(1), 89–109. Retrieved from http://www.sciencedirect.com/science/article/pii/S093336570100077X doi:10.1016/S0933-3657(01)00077-X PMID:11470218

Kukar, M., Kononenko, I., & Grošelj, C. (2011). Modern parameterization and explanation techniques in diagnostic decision support system: A case study in diagnostics of coronary artery disease. *Artificial Intelligence in Medicine*, *52*(2), 77–90. doi:10.1016/j.artmed.2011.04.009 PMID:21646000

Kumar, D., Sathyadevi, G., & Sivanesh, S. (2011). Decision Support System for Medical Diagnosis Using Data Mining. *International Journal of Computer Science*, *8*(3), 147–153. Retrieved from http://ijcsi.org/papers/IJCSI-8-3-1-147-153.pdf

Larrañaga, P., Calvo, B., Santana, R., Bielza, C., Galdiano, J., & Inza, I. … Robles, V. (2006). Machine learning in bioinformatics. *Briefings in Bioinformatics*. doi:10.1093/bib/bbk007

Lavrac, N. (1999). Selected techniques for data mining in medicine. *Artificial Intelligence in Medicine*, *16*(1), 3–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10225344 doi:10.1016/S0933-3657(98)00062-1 PMID:10225344

Lavrač, N. (1999). Machine learning for data mining in medicine. *Artificial Intelligence in Medicine*, 1999, 47–62. Retrieved from http://www.springerlink.com/index/gwvhwk72vwkpjaw9.pdf

Meyfroidt, G., Güiza, F., Ramon, J., & Bruynooghe, M. (2009, March). Machine learning techniques to examine large patient databases. *Best Practice & Research. Clinical Anaesthesiology*, *23*(1), 127–143. doi:10.1016/j.bpa.2008.09.003 PMID:19449621

Prompramote, S., Chen, Y., & Chen, Y. P. (2005). 5 Machine Learning in Bioinformatics. Machine Learning in Bioinformatics. In Bioinformatics Technologies (pp. 117–153). Springer Berlin Heidelberg.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(1), 3. doi:10.1186/2047-2501-2-3 PMID:25825667

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Russo, M., Ferrari, A., & Webb, C. (2012). *Microsoft SQL Server 2012 Analysis Services: The BISM Tabular Model*. doi:10.1016/S1754-4548(07)70038-4

SAS. (2014). Hadoop: What it is and why it matters. Retrieved from http://www.sas.com/es_co/insights/big-data/hadoop.html

SAS. (2014). Enterprise Miner 13.2. Retrieved from https://www.sas.com/content/dam/SAS/en_us/doc/factsheet/sas-enterprise-miner-101369.pdf

Silvia, P., & Claps, M. (2013). Bigger Data for Better Healthcare. *IDC Health Insights*, September, 1–24.

Soni, J., & Ansari, U. (2011). Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction. *International Journal (Toronto, Ont.)*, *17*(8), 43–48. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=09758887&AN=74184303&h=KvSpngkTsPtbMPAj4I04aBMcE1j6ncmWWYXI6hOisilndCDh+okd4iogsITZm98dgyGvOs6SWp3bjXiVImPjPg==&crl=c

Tamayo, P. (2005). Oracle Data Mining. In *Data Mining and Knowledge Discovery Handbook* (pp. 1315–1329). doi:10.1007/0-387-25465-X_63

Tomar, D., & Agarwal, S. (2013). A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio-Technology*, *5*(5), 241–266. doi:10.14257/ijbsbt.2013.5.5.25

Trifirò, G., Pariente, A., Coloma, P. M., Kors, J. A., Polimeni, G., Miremont-Salamé, G., ... Fourrier-Reglat, A. (2009). Data mining on electronic health record databases for signal detection in pharmacovigilance: Which events to monitor? *Pharmacoepidemiology and Drug Safety*, *18*(12), 1176–1184. doi:10.1002/pds.1836 PMID:19757412

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining*. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C

Yang, Q., & Wu, X. (2006). 10 Challenging Problems in Data Mining Research. *International Journal of Information Technology & Decision Making*, *05*(04), 597–604. doi:10.1142/S0219622006002258

KEY TERMS AND DEFINITIONS

Big Data: Large amount of data, in many different formats, that flows rapidly in real time. This data should undergo some sort of analysis in order to extract useful information or knowledge.

Cloud Computing: High performance computing service which does not require location setting, but instead use a grid of computers over the internet.

Data Mining: The process of extracting new, useful, understandable and previously unknown knowledge from information that might help in decision making.

Electronic Health (Medical) Record: The narrative record written by the nurse or the physician during patient examination.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hadoop: Hadoop is an open-source software framework from IBM, capable for storing and processing big data in a distributed fashion on large clusters.

Machine Learning: The subfield of artificial intelligence that uses learning algorithms in handling problems.

Medical Diagnosis: Investigating the symptoms and causes of certain disease, either by oral examination or laboratory tests.

This research was previously published in the Handbook of Research on Healthcare Administration and Management edited by Nilmini Wickramasinghe; pages 47-58; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 100
Wireless Body Area Network for Healthcare Applications

Danda B. Rawat
Georgia Southern University, USA

Sylvia Bhattacharya
Georgia Southern University, USA

ABSTRACT

Wireless Body Area Network (WBAN) is an emerging field of research which has been progressing rapidly in recent years. WBAN is a network utilized for continuous monitoring of physiological state of the subject, where the patient can perform his regular activities while his body parameters get measured continuously and are accessed by the physician remotely. This chapter provides a thorough survey of current WBAN technologies in the healthcare sector. Besides the recording of physiological parameters, discussions have been provided on remote data transmission to a server called Virtual Doctor Server (VDS). During this transmission, WBAN network uses various technologies namely Ultra Wide Band WBAN, Technology Enabled Medical Precision Observation 3.1 (TEMPO 3.1), J2ME and Bluetooth. Details of several existing WBAN related projects have been discussed along with their applications. The next section of the chapter deals with the use and design of medical sensors in WBAN. Performance comparison between WBAN and WSN (Wireless Sensor Network) has also been provided.

1. INTRODUCTION

WBAN is the joint application of biomedical science and wireless communication systems. The development of this technology will have the impact of drastically reducing the time a patient has to spend in his house or in a medical care center for treatment. This technology utilizes the principles of wireless communication to relay data (such as heart rate, ECG) continuously to remote areas (Mundt et. Al., 2005). The data are measured by sensors in a person's body. Besides drastic development in medical instrumentation, it is now also important to have compact machines like portable blood pressure measurement kits that can be attached to a body and be able to measure the parameters round the clock.

DOI: 10.4018/978-1-7998-1204-3.ch100

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

WBAN has a wide range of applications in medical science. Gait analysis, asthma detection, cancer detection, physiotherapy, brain activity analysis, etc. are some of the applications. In these applications, sensors are installed either inside or outside the body to determine the pattern of movement, to compare the result with standard data and to keep records of the physiological parameters of the patient. In this paper, it has been explained thoroughly how each of these physiological parameters are measured with the help of WBAN round the clock.

The recorded data are sent to a server where they are analyzed. These servers are the point from where the physician get access to the report of the patient round the clock. This server is called the Virtual Doctor Server (VDS). VDS performs a number of important responsibilities which includes maintaining records of patient data to build up the history of the patient, keeping track of patients' medicine timing and giving alarm to the patient to take the pills on time, giving advice to the patients assistant according to the physiological performance of the patient's body recorded by the sensors, alerting doctor or calling ambulance in time of emergency, maintaining confidentiality and privacy of patients' data records (Sghaier, et al., 2011).Patient can input his clinical feelings to the VDS system and then the VDS can give instant suggestion or advice based on the patients' history and vital signs. During this transmission, WBAN network uses various technologies namely Ultra Wide Band WBAN, Technology Enabled Medical Precision Observation 3.1 (TEMPO 3.1), ZigBee, J2ME and Bluetooth. Details of several existing WBAN related projects have been discussed in this paper along with their applications. The projects that have been discussed are Life Guard, Code Blue, Medisn, Wriscare, Mob health, WiMoCA, and Care Net (He et. al., 2013).

Biomedical sensors form a huge part of the patient monitoring system. A biomedical sensor is such an electronic device that senses, processes, sends or receives biomedical data. It has three main parts, the sensing unit, the processing unit and the receiving unit. There are micro controller units at the sensor node and a memory (RAM, ROM) to store the data (Lee et. al., 2006).The function of the sensor is controlled by software consisting of operating system. There are large varieties of biomedical sensors having separate functionalities such as electrocardiogram, electroencephalograph, etc. These sensors can be classified into in-body sensors and on-body sensors. Characteristics of each type of sensors are discussed in detail in this paper. Sensor networks have also been discussed extensively. They comprise of a large number of sensors and sink nodes and all of them are connected to each other wirelessly with the help of multi-hop or single hop communication systems. The data flows from the sensor node to the sink node which finally communicates the data to a gateway. The capabilities of the nodes vary with size, cost, battery lifetime, etc.

The security of WBAN is extremely important since it contains recorded data of patients and mishandling of such data may even cause death of the patient. The main factors affecting data security are confidentiality, dynamic integrity assurance, dependability. Security requirement is applicable to each device or communication link as data integrity, data confidentiality, authentication, availability, privacy, access control and non-repudiation. Each of these factors have been discussed in detail in this paper.

The remaining sections are organized as follows. Section II gives a brief idea about the background of WBAN along with a complete description of the process of WBAN. Section III describes the virtual doctor server which is the master server that works on the results obtained from patients body. Section IV describes the various technologies that are used in WBAN and the various ongoing projects in WBAN. Section V describes the sensors used in WBAN followed by a table showing the differences between the general sensor network and the wireless body network. Section VI is very important as it deals with

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the privacy and security of the whole system and at the end of this section the requirements for security in this type of network is stated. Section VII concludes the paper and presents the future goals in the commercialization of WBAN.

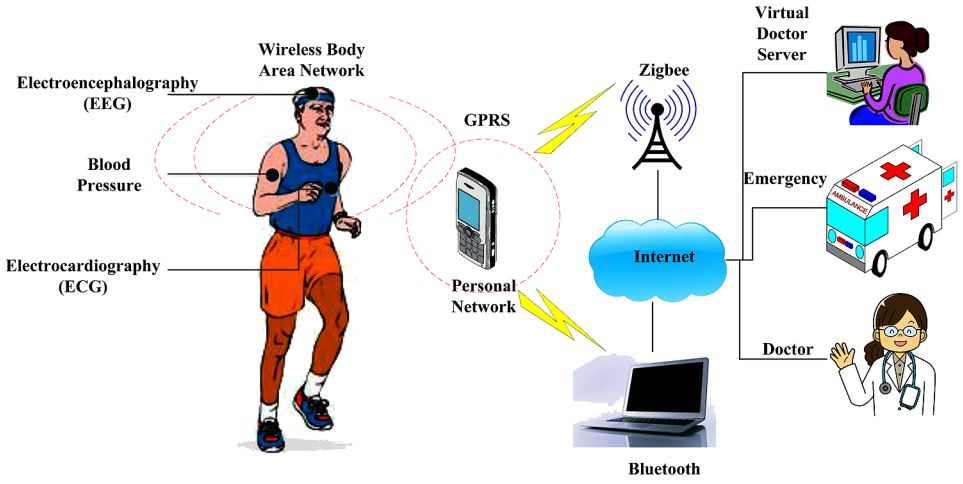
Figure 1 shows the whole mechanism of body area network. It shows how the personal area network records the physiological signals from the various sensors attached to the body and transmits the information to the virtual remote server or physician or emergency center.

The WBAN uses different wireless networks and access technology that leads to complex network as shown in Figure 1. One can easily see the difference in terms of data rate and average power consumption for WBAN applications. We note that Wireless USB and Wi-Fi based technologies offer high data rate but consume high power as shown in Figure 2.

2. BACKGROUND

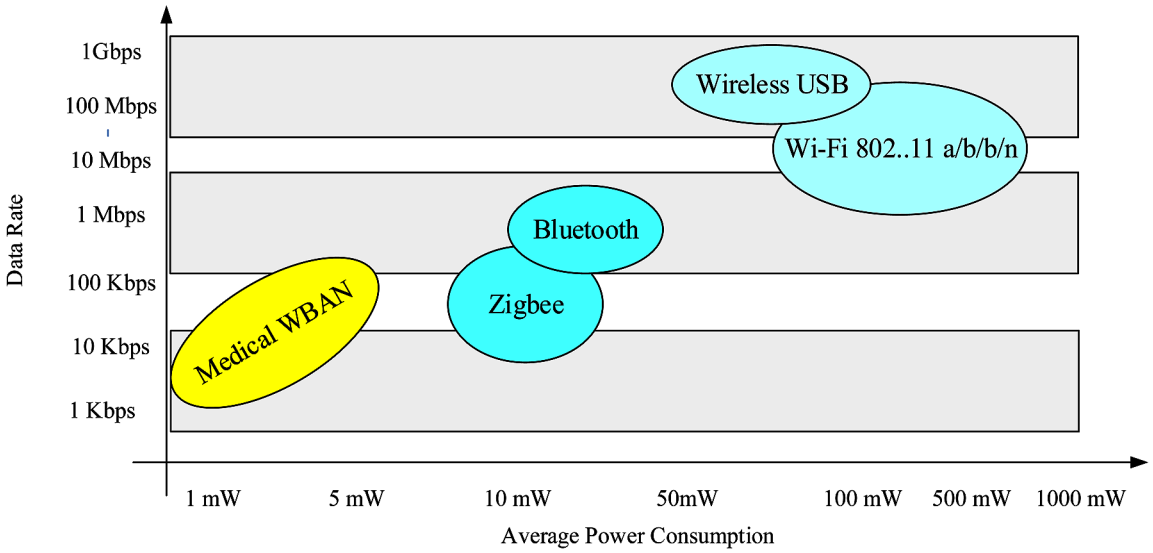
The technology of WBAN started developing in 1995 from the concept of Wireless Personal Area Network (WPAN) around the human body (Gafurov et. al., 2010). WBAN uses the technology of Wireless Personal area network (WPAN) as the gateway to communicate to longer ranges. With the help of this gateway, connection can be established with the human body sensors through internet and doctors or physicians can access the medical data of patient through this internet to keep the track round the clock. The drastic development in wireless technology and the increasing amount of sophistication in sensor technology has helped in the realization of the concept of WBAN. The medical sensors are used to monitor the medical parameters of the patients during treatment of chronic diseases such as asthma, blood pressure, diabetes etc. (Ivanov et. al., 2012) and to control patient rehabilitation (Bajcsy, 2007). The sensors have to be so designed such that they are comfortable and do not cause any harmful impacts on the body. An external processor transfers the information further after the sensing mechanism of the sensor is completed (Motoi et. al. 2006). Doctors who have the authority to access the data can visualize

Figure 1. Schematic diagram of a wireless body area network



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Data rate vs. power profile of different wireless technologies



the result from any part of the world. The doctor can send messages to the patient instantly through this system if he figures out any abnormalities or in emergency situations. Therefore, WBAN can be used to alert the hospital that a patient may experience a heart attack soon by measuring changes in their vital signs. WBAN can also be programmed to perform certain tasks in case of emergency conditions. For example, insulin can be injected in the body when blood sugar levels drops out. Other applications of WBAN can be extended to include sports, military and security services.

3. VIRTUAL DOCTOR SERVER (VDS)

VDS is an intelligent system that is designed in such a way that it can assist a physician in diagnosis of the medical parameters that are functioning abnormally. VDS also collects medical information about the patient from the medical history available from the doctor in order to understand and better diagnose the problem of the patient. They have a virtual doctor construction like a robot which mimics the physician. This robotic doctor interacts with the patient through their voice and examines the various physiological parameters of the patient in order to detect the problem in the health of the patient (Smith et. al., 2013). Diagnosis comprises of both physical and mental state of the patient namely physical ontology and mental ontology respectively. These two ontologisms are compared with each other in order to find similarity with medical history to diagnose the problem. These two ontologisms are compared with the help of Bayesian network which is based on the knowledge of expert doctors. This methodology has been verified by many researchers in their experiments and is successfully done. An integrated computerized model is constructed which mimics a human physician and with the help of it; an interface between that model and the real human user (patient) is utilized for initial diagnosis. Numerous research is already going on the field of WBAN. Apart from the general model of the WBAN, a VDS system is also added to it for better patient care after the physiological states are recorded by the sensors placed on the patient

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

body (Anis et. al., 2010). VDS is a smart intelligent system which records the physiological factors of the patient round the clock and sends the reports to the physician and in case of emergency can call the physician, relatives or even can give basic instructions to the patient.

The system consists of sensors both in body and on body placed in order to record physiological parameters like ECG, EEG etc. The data recorded by the sensors is then sent to a central processor unit in order to analyze and further process the recorded raw data. The mobile phone of the patient can also be used to send the data to the VDS system for further assessment and expert advice. This whole network system is named WBAN where each and every body is assigned an IP address to get distinguished from each other.

The interaction between patient and computer interface is stated in many paper which gives us further detail of the process. Two types of states are described in the research. The first one is the mental state which is the mental behavior of the patient due to a disorder and second is the physical behavior of the patient due to the disorder which is generally the consequence of such a disorder. The two types of bodily states are mapped and described using two software named OWL-S and SWRL which diagnoses the reason to such bodily state. An integrated computerized model is constructed which reflects a diagnostic system for human being as computer model and with the help of it, the patient and the system can interact with each other in order to diagnose the disorder caused to the patient. The VDS system use the knowledge that is obtained from UMLS for testing, and the integrated mapping of the two states of body which is represented through OWL-S framework.

The VDS system performs a variety of responsibilities which are as follows:

- Keeping record of patient data to build up the history of the patient.
- Keeping record of patients medicine timing and giving alarm to the patient to take the pills on time.
- Giving advice to the patients assistant according to the physiological performance of the patient's body recorded by the sensors.
- Calling doctor or ambulance in time of emergency.
- Maintain confidentiality and privacy of patients data records.
- Patient can input his clinical feelings to the VDS system and then the VDS can give instant suggestion or advise based on the patients history and vital signs.

Hence, the VDS is a very important part of the WBAN in terms of data analysis and for providing suggestions based on the medical data. On the other hand when patient is moving from one place to another, it is very important to have this kind of automatic intelligent system to take care of the patient's health giving him important suggestions and reminders.

4. WBAN TECHNOLOGIES

The different existing WBAN Technologies are Ultra Wideband WBAN, TEMPO 3.1, J2ME and Bluetooth. Each of these technologies are discussed in detail in this section.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

4.1 Ultra Wideband WBAN (UWWBAN)

Ultra WBAN is used for medical sensing and for in-body tracking and imaging (Khaleghi et. al., 2010). There is an antennae designed for on-body tracking that can communicate with same efficiency as in free space when the distance between the antennae and human body is 16mm (Anis et. al., 2010).

Primary function of UWWBAN is to interface between medical sensors and personal server (PS) in telemedicine systems (Francisco e.t al., 2009). UWWBAN can reduce the probability of detection for protecting patient's data with high security. This can reduce the need to hide the algorithms in sensors which reduces complexity and makes the device smaller in size. Hence the algorithms in the machine are required to be hidden and hence it gives a much smaller handy shape and size. UWWBAN does not interfere with any nearby signals and hence protects security and privacy of patients' data. UWWBAN has no tendency to react with any surrounding network and hence data remains completely protected within the system (Monton et al., 2008). This technology is mostly useful in future for transmitting information over short distance range with a very high speed. The bandwidth of this type is in the range of microwave and it is limited by the Federal Government rules (Takacs, 2006). The UW WBAN transmits data at very low rate pulse and hence it is very suitable for WBAN.

Ultra wide band (USB) radar is designed on the basis of impulse radio and hence it can detect the physiological parameters like blood pressure, insulin level etc. With the development of semiconductors the sensors of UWWBAN are made of low power consumption materials and hence are very efficient for wireless transmission task in WBAN. All the antennas in the system is simulated with the help of a software named Antennae a Simulator (Degli-Esposti, 2014). From all the research literature survey on UWWBAN, the most interesting application of UWWBAN is in Capsule Endoscopy. In this technology a camera is used which is small equivalent to a capsule size and is used to visualize gastrointestinal tract. They are designed in order to record images of digestive tract in order to detect gastrointestinal diseases .Even real time imaging of the gastrointestinal tract is possible (Fabre et. al., 1992).

Depending on their location of operation, there are three sensors associated with UWWBAN namely in- body sensor, on-body sensor and an external sensor (Alvarez-Folgueiras et. al., 2011).The in-body sensors are implanted inside the human body and on body sensors are placed on human skin whereas the external sensor is not in contact at all with the patient skin. Endoscope capsules are the in body sensors and the radars used are the external one. Once their inter connection is established they constitute the WBAN. In-body sensors are very critical and have the hardest structure. Hence an implanted device should be handled carefully as complexity of its circuit and its transmitting power is also very critical. The on body sensors are called gate nodes as their main function is to transmit information from the in body sensors to the WBAN. The circuitry of on body sensor is less complicated and most of its application is seen in Electromyography (EMG). Real time analysis of EMG is not necessary as in case of ECG. The principal component which detects and records all the medical information gathered by the sensors in WBAN is called network coordinator (NC). ECM-368 is the network coordinator which is used for the connection between the NC and P (Rasmussen et. al., 1996). It gives supports to data rate which is very high and also cognitive radio in some of the parts of Europe.

4.2 TEMPO 3.1

TEMPO 3.1 is an acronym for Technology Enabled medical Precision Observation 3.1. It is a third generation sensor which accurately captures and processes six degree of inertial data in a noninvasive

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

factor. The complete system of TEMPO comprises of Hardware and Software part design of TEMPO 3.1, a custom case design and an operating system that manages signal acquisition, processing and management of all data received. The system is then compared to BASN hardware platform (Wu, 2010).

BASN has its application in a number of fields namely healthcare, entertainment, fitness, etc. (Sanchez et. al. 2011). Recent advancement is BASN technology includes noninvasive technology to sense physiological signals and to transmit them across the body by low radio power. BASN should be designed in smaller size with smaller batteries. Many BASN sensors are now being developed for commercial and research use (Vos et. al., 2014). Hardware of BASN is now capable of real time imaging, have lithium batteries, have reduced size of the setup and developed sensor interfacing.

TEMPO 3.1 provides full rotational and translational sensing at the same site. The angular rate is measured by two separate MEMS gyroscope. Functions of TEMPO are data collection, streaming and signal processing. Microcontroller is also used for long term and which is energy efficient. The two networking protocols are Bluetooth and ZigBee which operate in 2.4GHz.

TEMPOS are real time operating system and it carries out signal acquisition, processing and control. They operate under many constraints like such as real time execution, memory utilization and minimum user input. It operates with the help of MSP430 microcontroller and for more efficiency allows access to low level hardware. For processing complex signals new signal processing algorithms are treated as application like Haar wavelet decomposition (Degli-Esposti, 2014). TEMPO3.1 is wearable and non-invasive with plastic and custom casing to allow minimal spatial overhead. There is no hardware user interface in the TEMPO setup except a charger point at the top of the machine. No interface is kept for the following reasons:

- If there are exposed buttons then its body cannot be designed water resistant.
- When it is working on a body, user can impatiently change the buttons in order to change the mode of operation which is not desirable.
- Users do not have to change their habit of wearing TEMPOS regularly as it has no buttons and can function autonomously.

On body sensors have already shown their use to a great extent. TEMPO 3.1 has certain impacts in several areas. There are lots of diseases that are cause due to human movements like Parkinson's disease, Gait disorder, cerebral palsy. Monitoring these diseases carefully can prevent, diagnose and treat these kinds of diseases. When a patient reports on his own by self-assessment, it does not give an accurate result and hence we should rely on sophisticated tools. Hence such self-reports are considered as under sampled, unnatural and unreliable. BASNs such as TEMPO 3.1 can correct such data by continuous monitoring all over the patient's body for a long period of time. Tremor suppressing deep brain stimulation and prosthetic control can be detected by motion capture feedback derived on body from TEMPO 3.1. TEMPO is now IRB approved for four clinical researches involving gait, tremor and akathisia. It has also its function in fitness. It is exclusively useful for athletes and amateurs alike. On body inertial sensing fills current niche which exists between pedometers and machine vision system.

4.3 J2ME and Bluetooth

With the continuation of research in wireless body area network, another technology came into prominence named Bluetooth area network which is operated by Bluetooth and Java technology. Smart phone

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

and Java are the central nodes of Bluetooth area network of a patient (Bratchikov, 2007). The general architecture of Bluetooth area network is as follows:

- A low power short range wireless interface medical sensor.
- Central coordinator communicates with sensors and it also performs the function of internet gateway to other networks in order to transmit bio signals from sensors and receive information.
- A central node which stores the sensor signals and detects alarm and sends patients physiological report to doctor via email, SMS, etc.

This technology is hence a monitoring network of Bluetooth biosensors which is connected to a 3G cell phone via WLAN. This setup is designed to perform the following functions:

- Visualizing and analysis of the physiological status of the patient.
- Recording the outdoor location of the patient using the system.
- Gives medical assistance in case of emergency.

There are three basic components of this Bluetooth system:

1. Internet Bluetooth Area Network (I BAN) to monitor the patient remotely,
2. A central control server (CCS), and
3. A mobile control monitoring unit which is carried by the doctor (MCMU. I BAN consists of a commercial pulse oximeter, a GPS and an Intelligent Node (IN).

Wi-Fi or GPRS is used to send information from IN to CCS and to MCMU (Fabre et. al., 1992).To keep the Bluetooth connection alive the device should be in coverage area of IN so that as soon as any emergency occurs the IN can inform the server. The data rate depends on the mode of operation of the BAN. There are two modes, namely Default mode (mode 1) and verbose mode (mode 2). When a MCMU receives an SMS alert, the physician should start the J2ME monitoring application which displays the patient and sensor data. The Java application is incorporated in the system for the IN. It uses the following application program interface (API):

- Bluetooth API JSR-82 which manages the Bluetooth.
- Location API JSR-179 to get more reliable data from GPS.
- Wireless messaging API JSR-205 for sending SMS and MMS.
- File connection API JSR-75 which records patient's data in mobile phone and memory chip.

The MMS sent by the IN is in the form of ordinary MMS which gets directly into the inbox of the cell phone. The doctor can access the network from any place through internet using Remote Method Invocation which is supported by Java application (Alvarez-Folgueiras et. al., 2011).

Some of the popular WBAN projects being undertaken around the world are listed below:

- Lifeguard (Mundt et. Al., 2005) is a set up that was designed for the astronauts and can be used for monitoring vital signals from patient. It is capable of monitoring various kinds of data such as heart rate, blood pressure, respiratory and pulse oximeter. The physiological data are collected by

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

a wearable device called Crew Physiological Observation device (COPD). It is comprised of skin temperature sensors and 3 axis accelerometer. There is a Bluetooth tablet that displays the data and stores it from COPD for further future analysis.

- Code Blue (Sghaier, et al., 2011) is another hardware and software platform that is designed by Harvard University. The hardware includes a design for 2 lead ECG system, pulse oximetry and a board for motion analysis. The software architecture is based on a routing framework and each end user communicates with this Code Blue system by a specialized program named Pub Subbase.
- Medisn (Lee et. al., 2006) is another example of a set up that monitors the unattended patients in emergency rooms of John Hopkins Hospital. The set up has four parts namely mi Tag, mi Net, mi Store and mi Viewan. mi Tag consists of all the medical information tags and mi Net is the wireless network connection for communicating the in formations from mi Tag. mi Store stores all the in formations regarding the collected patient data and mi View an is used to display the data to some end users who are designated to analysis the patient data.
- WrisCare (Kim et. al., 2009) is a setup for monitoring heart rhythms for a long period of time. It can detect the changes taking place with heart rhythms. It is of two types namely, home system and institutional system. The home system consists of a wrist band and a base station and is used by the people who live independently in their own houses while the institutional system is used by people who are admitted in Hospitals or Nursing home.

5. SENSORS AND WBANs

Any wireless network comprising of sensors connected to each other is called Wireless Sensor Network (WSN). Basically, in WBAN the sensors are of three types, in-body sensor, on-body sensor and network conductor. This wireless communication system relieves from the pain that is caused due to the cables induced in wired medical communication system. Body area Sensor network nodes should have ergonomic design, have small factors and smaller batteries. A third generation body area sensor is also designed called TEMPO (Technology Enabled Medical Precision Observation) that accurately captures data of human movement and transmits it wirelessly. They can process the data continuously as it is a real time operating system that manages signal processing, control acquisition and everything (Shih-Yeh Chen et. al., 2011). In some wireless area networks mobile phone is used for collecting data. A lot of research is done by the scientists where they used android phone as the medium to communicate the whole process (He et. al., 2013). A light weight activity recognition is detected due to data constraint.

WSN is composed of thousands of nodes and each node is connected to each other as a network so that information gets transmitted from one node to the other (Wu, 2010). These nodes again have several parts. They are microcontroller, a circuit that connects with both the sensor and an energy source, a radio transceiver which is either connected to an external antennae or an internal one, and a battery or any other energy source device. The size of sensor node varies from one another. Similarly the cost of sensor nodes also varies depending on the complexity of the node. Depending on the size and cost constraints of sensor nodes resource constraints such as energy, memory, computational speed and communications bandwidth depends. Routing is the propagation technique between the hops of the networks. WSN topology varies from star network to mesh network which advanced and multi hop wireless.

Biomedical sensors are used mainly in three types of places (Rasmussen et. al., 1996). They are hospital, nursing home and patient's residence. Progressive work is done in this field and a lot of demonstrators

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

are found. The demonstrator for emergency is Field care demonstrator, WiSMoS demonstrator is for post-operative monitoring whereas blood glucose demonstrator is mainly used for monitoring patients at home (Polastre et. al., 2005). The working principle of all these demonstrators are same. They use biomedical sensor to collect patient data and the data is sent to a data collector for further processing and then that data is further sent to the hospital monitoring system for final analysis. The person who is responsible for processing and analyzing these data can access them directly. The principle properties of a WSN are as follows:

- Power consumption constraints for nodes using batteries or energy harvesting.
- Ability to cope with node failures (resilience).
- Mobility of nodes.
- Heterogeneity of nodes.
- Scalability to large scale of deployment.
- Ability to withstand harsh environmental conditions.
- Ease of use.
- Cross-layer design.

Cross-layer is becoming an important studying area for wireless communications.

WBAN and WSN are part of each other and have many things similar to each other. But the basic differences of the two network systems are stated in Table 1.

6. SECURITY AND PRIVACY

Not much work has yet been done regarding the privacy and security of the system. Data security means securing a database from unwanted actions of unauthorized users. It is the relation between collection and dissemination of data, technology, privacy and legal and political issues around. While starting with security of WBAN, we should keep in mind the efficiency of the system too. Even if the device is compromised the patient related data should always remain confidential. Medical data is very vital and it will result into disastrous result if it is modified. The three main factors that makes it important for keeping the data secured are:

Table 1. Comparison of WSN and WBAN

Factor	WSN	WBAN
Subject	Environment	Human Body
Required Security Level	Low	Higher
Bio Compatibility	Not required	A must
Mobility	Stationery	Moves from one place to another
Data rate	Event based	Continuous data
Power	High	Low
Node Size	Large or Small	Essentially small

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- **Confidentiality:** The data that is recorded constantly from the patient's body is stored for the consultancy of the physician but it needs to be kept at a secured location in order to prevent unwanted people from accessing them. The data is usually stored at a node or a local server. Here the data does not get leaked during their storage time. The data confidentiality must be resilient to device compromise to attack.
- **Dynamic Integrity Assurance:** A patient who has WBAN for detecting his physiological parameters are critical patients and hence their data are very vital. A little bit of modification of the original data might lead to dangerous consequences. Hence data integrity shall be protecting dynamically all the time (He et. al., 2006). This will help to detect the changes made by the end user and also check that during storage periods in order to detect malicious modifications and an alert is sent to the user.
- **Dependability:** WBAN data are very important data and should be retrieved any time. If there is any problem with getting data on time it can be life threatening. Fault tolerance is very useful in order to manage the threats caused by network dynamics. Hence patient data should be managed such a way that it can retrieved anytime even during node failure or some malicious modifications (Huang et al 2013).

Authentication is a very important step for securing data. It prevents false data injection, DoS attacks and also it verifies the identity of the user before letting the person to access the data. Moreover it is very essential to secure data during transfer with in WBAN system. When a person's data is recorded and processed and transferred from one station to another for further analysis, it is handled by more than one user and hence this path is very risky. Any user can hack the data and modify it. So, it is very important to have user authentication at each step of data transfer where user needs to put a password and get into the system to access the actual data (Xuange et. al., 2010).

High efficiency is very important for the security of a WBAN. The sensors used in recording the parameters should be light and efficient in design and if the authentication protocol is not fast enough then DoS can attack easily (Sung-Yuan, 2001). Privacy of the patient data is destroyed if it is accessed publicly and hence security is one of the most important thing of WBAN. It is very difficult to maintain privacy and security of the data together. Hence it is very difficult to design such a system which is practical. Very few studies are done on this and even the one those are done are far from mature. In Code Blue project a sensor network is developed for pre hospital and emergency responses (Sanchez et. al. 2011). In this system an epileptic curve cryptographic based public key is used for authentication. But there is no other security for stored data and even the access cannot be controlled.

Chess a et al. proposed Redundant Residue Number System which is a secured data storage and sharing scheme for wireless mobile networks (Vos et. al., 2014). In Wang's scheme privacy, dynamic integrity and dependability is achieved all together which is important they usually contradict which each other. Pietro et al found out the solution for data survival problem in wireless (Atallah et. al., 2009). Fine-grained distributed control is an important security service in healthcare. It is designed as role based access model. It prevents handling the patient data publicly and restricts its users with authentication according to their roles. There are 3 types of schemes that are proposed namely SKC based, and ABE based (Ivanov et. al., 2012). SKC was proposed by Morchon et al. (2015). In this method patient can establish connection easily with one entity after authentication. But there are certain disadvantages in the SKC scheme. They are as follows:

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

- Due to high key management complexity, fine gain access control is hard to be achieved.
- User collision can take place.
- Compromising a node may leak the data or information.

ABE is a very effective scheme. Here the cipher text is allowed to be read by only a group of user who satisfy a certain authentication process. It is collision free. Whenever any data is generated by a WBAN an authentication verification goes with the cipher text (Bhoir et. al., 2014), it supports tree like access policy structure and is easily expressive and it is easy to integrate the context related parameters as attributes like time. One point multiplication and pairing operation takes seconds and hence it might not be a good choice on the sensor nodes (Kathuria et. al., 2014). But it is capable than any other technique as far as obtained after research .but none of the schemes satisfactorily support the security and hence further research in this field is required.

Bluetooth is another extremely popular technology which is short range and low power which is used for portable computing and communication devices. Bluetooth establishes connection automatically which is a great advantage but it has some risks too. There is a huge chance of data getting vulnerable with other low power radio waves (Kathuria et. al., 2014). Unwanted people may receive private information and even people can send files or virus without any permission. The weakest part of the blue tooth technology is that, while establishing connection for the first time between two devices, an authentication is done by each other but without encryption and hence if the hacker actually gets to know the password then he can easily guess the initiation keys and can get all the data that is transferred between the two devices. A very common way of attacking through blue tooth is during pairing process. At this time the connection between the devices is forcefully broken by flooding the channels with packets indicating the slave has lost the key. And when the pairing process is repeated, the hacker observes the encryption keys and gets the link. The BD ADDR is globally unique just like the MAC address (Shen et. al., 2013). In 2005, the researchers at CSAIL found that there is a big vulnerability in it (Ramachandran et. al., 2007). An attacker can assemble list of commonly used BD ADDR in a device and can change his Blue tooth module setting and starts getting packets .then he can attack by sniffing packets or injecting packets. Integrity threat and denial of service are very common in blue tooth communication. Some of the reported attacks in Bluetooth are:

- MAC spoofing attack.
- Attack to crack pin.
- Blue jacking attack.
- Blue printing attack.
- Blue snarfing attack.
- Blue bugging attack.
- Blue over attack.
- Cabir attack.
- DOS attack.
- Skull worms attack.
- Lasco worm attack.
- Brute force attack.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

GSM is a standard network used worldwide but it is also attacked. GSM uses several cryptographic algorithms are used by GSM for security. Air voice privacy is protected by using A5/1 and A5/2 stream ciphers. Denial of Service attacks are one of the most powerful attacks. They are so dangerous that they hack the patients' medical data and substitute those valuable data with insensible bogus messages. Now a days a new scheme of location signature is introduced. Symmetric or asymmetrical key cryptography is used which encrypts / decrepit to hide the medical data of patient. Biometric encryption is widely used in most of the places but it is very weak. No, method is strong enough to protect the medical data of patient, so, the security should be strengthened for WBAN to be successfully commercialized. The security requirements of each device and communication link are listed below:

- **Data Integrity:** It assures the correctness of data protection against modification, deletion or replication from an unauthorized user. Usually MAC provides data integrity using secret keys.
- **Data Confidentiality:** Encryption and access control links are the main methods that ensure data confidentiality. It protects the data from handling by unauthorized users. Medical data is very vital and hence it is very important to keep it confidential.
- **Authentication:** This is process when a verification is done with a password in order to choose the right user. There are two kinds of authentication namely message authentication and entity authentication. Biometrics is the most trustworthy authentication possible. There is an ID certificate for the user and a device certificate for the WBAN which achieve authentication.
- **Availability:** This protects from the event taking place and impacting the network. There can be many existing attacks and they results in loss of availability. There are some physical actions that are taken for availability in order to make up for the loss of availability.
- **Privacy:** The medical data may consists of some ones routine health checkup or a particular physiological parameter .These data need some privacy and should not be shared with everyone. Privacy is very important for a patient and these data can only be accessed by people who are allowed to do so.
- **Access Control:** As firewall provides access control, there is personal server that works as a part of firewall. This protects from unauthorized user to access the medical data. It remains open to only few people who are part of WBAN to access data of patient.
- **Non Repudiation:** The process which always keeps evidence. When a user sends or receives it always keeps an evidence of the same. This evidence can be presented to a third party in order to prove that such event has taken place in case of any problem.

7. CONCLUSION AND FUTURE GOALS

Wireless body area network (WBAN) is an emerging field of research which has been progressing rapidly in recent years. The development of this technology will have the impact of drastically reducing the time a patient has to spend in his house or in a medical care center for treatment. This technology has numerous advantages and is very safe helpful if proper security system is incorporated. This technology can get even more popular in use once its privacy and security section is built high. With this development, it can be one of the most successful innovations in medical science. The security issue of this system is very important and from the survey that is performed it can be seen that there is awareness of the importance of these patients' data but not sufficient work has been done yet on this topic. Many scientists have

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

built some models to protect a part of the system. But, the patient monitoring system consists of several process and each and every step should be highly secured so that the information doesn't get leaked at any point of the network. The main focus of this paper has been to draw attention of the researchers to this technology to develop it to make it more commercialized. The future goal includes making it more cheap, safe and popular.

REFERENCES

Alvarez-Folgueiras, M., Minana-Maiques, M., Moreno-Piquero, E., Jorge-Barreiro, F., Lopez-Martin, E., & Ares-Pena, F. (2011, April). Experimental system forthe study of multi-frequency dosimetry. In *Proceedings of the 5th european conference on antennas and propagation* (pp. 1296-1299).

Anis, M., Ortmanns, M., & Wehn, N. (2010, May). Fully integrated uwb impulse transmitter and 402-to-405mhz super-regenerative receiver for medical implant devices. In *Proceedings of 2010 ieee international symposium on circuits and systems* (pp. 1213-1215). 10.1109/ISCAS.2010.5537294

Atallah, L., Lo, B., Ali, R., King, R., & Yang, G.-Z. (2009, November). Real-time activity classification using ambient and wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, *13*(6), 1031–1039. doi:10.1109/TITB.2009.2028575 PMID:19726267

Bhoir, S., & Vidhate, A. (2014, Jan). An improved wban mac protocol. In *International conference on computer communication and informatics* (pp. 1-6).

Bratchikov, A. (2007, Sept). Unified biophysical model for interaction of low intensity laser and ehf electromagnetic fields with cell biosystems and integral organizm. In *17th international Crimean conference microwave telecommunication technology* (pp. 774-775). 10.1109/CRMICO.2007.4368937

Chen, S.-Y., Lee, W.-T., Chao, H.-C., Huang, Y.-M., & Lai, C.-F. (2011, Nov). Adaptive reconstruction of humanmotion on wireless body sensor networks. In *First international conference on wireless communications and signal processing* (pp. 1-5).

de Francisco, R., Huang, L., & Dolmans, G. (2009, Sept). Coexistence of wban and wlan in medical environments. In *IEEE 70th vehicular technology conference fall* (pp. 1-5). 10.1109/VETECF.2009.5378807

De Vos, J., Flandre, D., & Bol, D. (2014, May). Switchedcapacitor dc/dc converters for empowering internet-of-things socs. In IEEE faible tension faible consummation (FTFC) (pp. 1-4).

Degli-Esposti, V. (2014, April). Ray tracing propagation modelling: Future prospects. In *8th European conference on antennas and propagation* (pp. 2232-2232).

Fabre, J., Camart, J., Prevost, B., Chive, M., & Sozanski, J. (1992, Oct). 915 mhz interstitial hyperthermia: Dosimetry from heating pattern reconstruction based on radiometric temperature measurements. In *14th annual international conference of the IEEE engineering in medicine and biology society* (Vol. 1, pp. 231-231).

Gafurov, D., Snekkenes, E., & Bours, P. (2010, April). Improved gait recognition performance using cycle matching. In *24th international conference on advanced information networking and applications workshops (waina)* (pp. 836-841). 10.1109/WAINA.2010.145

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

He, Z., Luo, Y., & Liang, G. (2013, April). Runking: A mobile social persuasion system for running exercise. In Computing, communications and it applications conference (pp. 74-78).

Huang, C.-W., Kuo, S.-W., & Chang, C.-J. (2013, July). Embedded 8-bit aes in wireless bluetooth application. In *International conference on system science and engineering (ICSSE)* (pp. 87-92). 10.1109/ICSSE.2013.6614638

Ivanov, S., Foley, C., Balasubramaniam, S., & Botvich, D. (2012a, November). Virtual groups for patient wban monitoring in medical environments. *IEEE Transactions on Bio-Medical Engineering*, *59*(11), 3238–3246. doi:10.1109/TBME.2012.2208110 PMID:22801487

Ivanov, S., Foley, C., Balasubramaniam, S., & Botvich, D. (2012b, Nov). Virtual groups for patient wban monitoring in medical environments. *IEEE Transactions on Biomedical Engineering, 59*(11), 3238-3246.

Kathuria, M., & Gambhir, S. (2014, Feb). Quality of service provisioning transport layer protocol for wban system. In *International conference on optimization, reliabilty, and information technology* (pp. 222-228). 10.1109/ICROIT.2014.6798318

Khaleghi, A., Chavez-Santiago, R., & Balasingham, I. (2010, October). Ultra-wideband pulse-based data communications for medical implants. *Communications, IET*, *4*(15), 1889–1897. doi:10.1049/iet-com.2009.0692

Kim, K., Lee, I.-S., Yoon, M., Kim, J., Lee, H., & Han, K. (2009, Dec). An efficient routing protocol based on position information in mobile wireless body area sensor networks. In *First international conference on networks and communications* (pp. 396-399). 10.1109/NetCoM.2009.36

Kim, T.-Y., Youm, S., Jung, J.-J., & Kim, E.-J. (2015, Jan). Multi-hop wban construction for healthcare iot systems. In *International conference on platform technology and service* (pp. 27-28). 10.1109/PlatCon.2015.20

Lee, D.-S., Lee, Y.-D., Chung, W.-Y., & Myllyla, R. (2006, Oct). Vital sign monitoring system with life emergency event detection using wireless sensor network. In *5th IEEE conference on sensors* (pp. 518-521).

Monton, E., Hernandez, J., Blasco, J., Herve, T., Micallef, J., Grech, I., ... Traver, V. (2008, February). Body area network for wireless patient monitoring. *Communications, IET*, *2*(2), 215–222. doi:10.1049/iet-com:20070046

Motoi, K., Ikeda, K., Kuwae, Y., Yuji, T., Higashi, Y., Nogawa, M., ... Yamakoshi, K. (2006, Aug). Development of an ambulatory device for monitoring posture change and walking speed for use in rehabilitation. In *Proceedings of the 7th IEEE international conference on engineering in medicine and biology society* (pp. 5940-5943). 10.1109/IEMBS.2006.259364

Mundt, C., Montgomery, K., Udoh, U., Barker, V., Thonier, G., Tellier, A., ... Kovacs, G. (2005, September). A multiparameter wearable physiologic monitoring system for space and terrestrial applications. *IEEE Transactions on Information Technology in Biomedicine*, *9*(3), 382–391. doi:10.1109/TITB.2005.854509 PMID:16167692

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Polastre, J., Szewczyk, R., & Culler, D. (2005, April). Telos: enabling ultra-low power wireless research. In *Fourth international symposium on information processing in sensor networks,* (pp. 364-369). 10.1109/IPSN.2005.1440950

Ramachandran, A., Zhou, Z., & Huang, D. (2007, May). Computing cryptographic algorithms in portable and embedded devices. In IEEE international conference on portable information device (pp. 1-7). doi:10.1109/PORTABLE.2007.47

Rasmussen, J., Scholl, B., Gellekum, T., & Schmitt, H. J. (1996, Oct). Fiber-optic polarimetric temperature sensor for characterizing a 900 mhz tem cell used in bioeffects dosimetry studies. In *Proceedings of the 18th annual international conference of the IEEE engineering in medicine and biology society, bridging disciplines for biomedicine*. (Vol. 5, pp. 1875-1876). 10.1109/IEMBS.1996.646298

Sanchez, A., Blanc, S., Yuste, P., & Serrano, J. (2011, June). A low cost and high efficient acoustic modem for underwater sensor networks. In IEEE - Spain oceans (pp. 1-10). doi:10.1109/Oceans-Spain.2011.6003428

Sghaier, N., Mellouk, A., Augustin, B., Amirat, Y., Marty, J., Khoussa, M., . . . Zitouni, R. (2011, July). Wireless sensor networks for medical care services. In *7th international conference on wireless communications and mobile computing conference* (pp. 571-576). 10.1109/IWCMC.2011.5982596

Shen, Q., Liu, J., Yu, H., Ma, Z., Li, M., Shen, Z., & Chen, C. (2013, Aug). Adaptive cognitive enhanced platform for wban. In IEEE/CIC international conference on communications in China (pp. 739-744). doi:10.1109/ICCChina.2013.6671208

Smith, D., Miniutti, D., Lamahewa, T., & Hanlen, L. (2013, October). Propagation models for body-area networks: A survey and new outlook. *Antennas and Propagation Magazine, IEEE*, *55*(5), 97–117. doi:10.1109/MAP.2013.6735479

Sung-Yuan, K. (2001). The embedded bluetooth ccd camera. In Tencon 2001. Proceedings of IEEE region 10 international conference on electrical and electronic technology (Vol. 1, pp. 81-84). doi:10.1109/TENCON.2001.949556

Takacs, B. (2006, July). A portable ultrasound guidance and raining system using high fidelity virtual human models. In *International conference on medical information visualisation - biomedical visualisation* (pp. 77-81). 10.1109/MEDIVIS.2006.2

Wu, P. (2010, Nov). The perspective of biomedical electronics. In IEEE conference on sensors (pp. 1187-1187). doi:10.1109/ICSENS.2010.5690645

Xuange, P., & Ying, X. (2010, March). An embedded electric meter based on bluetooth data acquisition system. In *Second international workshop on education technology and computer science* (Vol. 1, pp. 667-670). 10.1109/ETCS.2010.624

This research was previously published in Advanced Methods for Complex Network Analysis edited by Natarajan Meghanathan; pages 343-358; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 101
Information and Communication Technologies in the Healthcare:
Future Trends for Project Success

Jorge Gomes
Universidade de Lisboa, Portugal

Mário Romão
Universidade de Lisboa, Portugal

ABSTRACT

Web portals, sensors, electronic records, video communication, on-line gaming and mobile apps. These are all examples of information and communication technologies (ICT) applications or devices that might cause benefits to healthcare. The ICT has the potential to dramatically change the way individuals or society see the sector, and provide tremendous opportunities for supporting professionals, improving effectiveness and efficiency. Healthcare organizations have become more and more challenged on how to assure a fair return from ICT investments. Thus, the application of project management in health is important because it allow more productivity and, as a direct result, more accessibility, higher quality care and a safer environment to patients. The study of the success of these initiatives has become vitally important for the hospitals performance. The article collects relevant data and provide recommendations about the perceived benefits of ICT project implementations, proposing a review of the published work to provide some insights into the benefits of these implementations.

DOI: 10.4018/978-1-7998-1204-3.ch101

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The world has witnessed significant health gains in recent decades due not only to advances in science, technology and medicine, but also in expanding infrastructure, rising incomes, better nutrition, sanitation, literacy and opportunities for women (Dzenowagis, 2005). While promoting population health has been the classic goal of public health practices and policy, in recent decades new objectives in terms of autonomy and equality have been introduced (Munthe, 2008).

Healthcare systems around Europe are facing major challenges related to chronic diseases, demographic changes, nursing shortages, medical accidents and increasing costs (EC, 2009; WHO, 2011, 2015). The proportion of people over 65 is expected to almost double by 2050 (Eurostat, 2005) and average public spending for health and long-term care in OECD countries may rise to 10-13% of the GDP (WHO, 2011). Today, the problem of improving healthcare delivery in the developed countries is more about the equitable distribution of available resources to the population than about technological issues (Bukachi & Pakenham-Walsh, 2007).

Since the 1990´s, the health sector has sought to improve its effectiveness and efficiency by adopting ICT solutions to increase the levels of services quality, namely, patient safety, organisational efficiency and patient satisfaction (Bates and Gawande, 2000). The aim of ICT for healthcare is to improve the quality, access and efficacy of healthcare for all citizens (Ammenwerth et al., 2004; EC, 2009). ICT for healthcare have been referred as a key instrument that facilitate communication, processing and transmission of information by electronic means, with the aim of improving health (Bukachi & Pakenham-Walsh, 2007; Drury, 2005; EC, 2009; Häyrinen et al., 2008). The ICT usage in health is a means to reach a series of desired outcomes across the entire health ecosystem and are part of the continuous quality improvement of healthcare: According to Brandrud et al. (2011), three main key success factors must be guaranteed: (1) Reliable information; (2) Engagement of all stakeholders in all phases of the work improvement; (3) A proper infrastructure involving multidisciplinary teams.

ICT in healthcare represents the integrated effort to collect, process, report and use health information and knowledge to influence policy-making, program action and research and further states that they are essential to the effective functioning of health systems worldwide (WHO, 2006). Furthermore, ICT underpin the expansion of the world's health knowledge: linking systems, enabling research, improving capabilities and communication for healthcare (EC, 2009).

METHODOLOGY

This study is aimed to research the pattern of ICT implementations in healthcare units, using articles published in academic health management journals. The documents were retrieved based on a literature review undertaken using the keywords 'ICT in Healthcare' and 'ICT in Hospitals', available in the following sources: ProQuest, Sciencedirect/Elsevier, Ebscohost, Emerald, Jstore, Medline/Pubmed, Taylor & Francis Online, Sagepub, and Wiley databases. The final set of selected articles resulted from the comparison between the documents collected and the objectives of our study.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

ICT Implementations

The introduction of ICT can radically affect the healthcare organizations and the healthcare delivery (Ammenwerth et al., 2004). Although the increasing role and importance of ICT in healthcare, the understanding of why, how, and when technology enables the productivity improvement is in its early stages, resulting most often in a lag in productivity growth (Jorgenson & Stiroh, 2000). The main goal of ICT is to manage information from all healthcare-related activities, including planning, monitoring, coordination, and decision making. The real-time access, exchange and receipt of clinical data provided by the system have improved clinical documentation, reduced the duplication of care services, and supported better decision making related to patient care (Mäenpää et al., 2009). It is commonly accepted that the use of ICT offers tremendous opportunities to support healthcare professionals and to increase the efficiency, effectiveness and appropriateness of care (Kohn et al., 2000).

Nowadays ICT implementations are mostly meant to extensively support disparate conventional data processing tasks. Ranging from simple application systems, such as transaction processing systems, to more complex ones, such as clinical decision support systems, here are some examples: (1) Patient centered system - manage comprehensive patient care information such as medical records and appointment scheduling (Cliff, 2012; Krist & Woolf, 2011; Snyder et al., 2011); (2) Administrative support - record the main business processes and routine transactions of organizations such as patient admission, discharge and transfer, or bill processing (Jiang et al., 2000; Peabody et al., 2004); (3) Clinical system - collection of specific data for patient care, research, management, planning and maintenance of national data repositories (Ammenwerth & de Keizer, 2005; Gardner & Shabot, 2001); (4) Specific systems (Radiology, Laboratory…) - support the acquisition and analysis of specific function of the different departments (Paré & Sicotte, 2001); (5) Telemedicine - provides and supports healthcare services across distances via electronic communications and IT (Gawande & Bates, 2000); (6) Clinical decision support systems - designed specifically to aid clinical decision making (Kaushal, et al., 2003); (7) Hospital information system - integrated hospital information processing systems. Support healthcare activities at the operational, tactical and strategic levels (Ammenwerth & de Keizer, 2005; Van der Meijden et al., 2003) (8) Cloud computing services such as software as a service application - explored by the scientific and medical communities for their potential in advancing and accelerating research and reducing costs (Sultan & Sultan, 2012).

The focus of investments in ICT projects lies mainly in improvements on the organizational performance and business efficiency, in other words, improving processes and changing the ways the work is performed (Ammenwerth, 2004; Ward & Daniel, 2012; WHO, 2011, 2015).

Seven different stages were proposed in the development of health information system (HIS) (Haux, 2006): (1) Shift from paper-based systems to computer-based processing and storage; (2) Shift from local to global information system architectures; (3) Health information system used by professionals and patients/consumers; (4) Data used for patient care, healthcare planning and clinical research; (5) Shift of focus from technical health IS problems to change management and strategic information management; (6) Shift from alpha-numerical data to clinical images and data on a molecular level; (7) Steady increase in new technologies for continuous monitoring of health status.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In various aspects, these implementations are different from project implementations in other industries. The diversity of systems and devices, together with the growing challenge of integration and interoperability are the main issues of the health environment (Abouzhara, 2011). The academy agrees with the existence of success-related factors for projects, namely the Project Success Criteria (PSC), referring to the group of principles or standards used to determine project success; and, the Critical Success Factors (CSFs), referring to specific conditions and circumstances that contribute to project success. While the potential and benefits from the use of ICT projects in health are large, the risks are also substantial. Therefore, it seems wise to advocate that organizations should give more attention to adopting formal project evaluations and benefits management methodologies to ensure the realization of the expected benefits from investments in those project implementations (Schultz, 2006; Ward & Daniel, 2012). Studies have shown significant correlation between organizational maturity and ICT project success, and project management practices and ICT project success (Gomes, 2018; Thomas & Mullaly, 2008), which reinforces the value of applying these tools.

ICT Project Maturity

Concerning the ICT environment, maturity is regarded as a measure to evaluate the capabilities of an organization (de Bruin, & Rosemann, 2005). Maturity models in ICT projects are understood as tools that facilitate internal and/or external benchmarking while also showcasing future improvement and providing guidelines through the evolutionary process of organizational development and growth (Mettler et al., 2010). Not surprisingly, the concept of maturity was also diffused to the project management field. Different views exist on maturity related to project management. These approaches assume that increases in project management maturity will lead to more consistent and more successful project outcomes. A significant number of competing perspectives attempt to describe mature project management practices (Cooke-Davies & Arzymanow, 2003; Ibbs & Kwak, 2000; Jachimowicz, 2003; Skulmoski, 2001).

Some studies have discussed the correlation between the level of project management maturity and project performance (Ibbs & Kwak, 2000; Jiang et al., 2004). The theory highlighted that investment in organizational maturity and in project management practices increases the results in enhanced project performance and project success that should translate into cost saving and other benefits (Gomes, 2018; Jiang et al., 2004). Due to their nature, maturity models are an important managerial tool that have been used extensively in ICT development (Wetering & Batenburg, 2009; Crawford, 2006). Studies supported the assumption that a positive correlation exists between maturity and project success. Higher levels of organizational maturity are expected to be more successful in terms of project effectiveness and efficiency and foster a competitive advantage in the marketplace (Cooke-Davies & Arzymanow, 2003; Gomes, 2018; Ibbs & Kwak, 2000; Sergeant et al., 2010; Skulmoski, 2001; Young et al., 2014).

ICT Project Failures

Despite remarkable technical progress, failures have still been reported when integrating technically sound systems into processes of care (Lorenzi & Riley, 2003). The investments on ICT in healthcare are financially relevant and still growing worldwide. Many results have revealed a waste of financial resources in acquiring large sized systems, which are totally ineffective (Heeks, 2006).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Studies have identified high failure rates in ICT projects in various sectors, particularly in hospitals. Heeks (2008) states that 35% of IS/IT projects are total failures and 50% partial failures, with only 15% being considered successful. A study by Gheorghiu (2006) found that 70% to 80% of all information technology and information systems fail. Similarly, Kaplan and Harris-Salamone (2009) confirmed a value greater than 30% for the failure rates of major health information technology projects. Project failures in healthcare are in part due to the following reasons (Lewis et al., 2011): (a) Lack of senior management commitment being incomplete or missing altogether (Bukachi and Pakenham-Walsh, 2007; Liu et al., 2010); (b) Difficulties in the engagement of health professionals, and a lack of focus on end-users (Elder and Clarke, 2007); (c) Incorrect specification requirements (Lucas, 2008; Gauld, 2007); (d) An absent or inadequate process of change (Yeo, 2002); (e) Poor knowledge of the complexity of health systems (Al-Ahmad et al., 2009); (f) Missing investments in human resources (Elder and Clarke, 2007; Bukachi and Pakenham-Walsh, 2007); (g) Proper training (Allen et al., 2000).

The study of the success or failure of these initiatives has become vitally important for the performance of health organisations. They must then improve their business practices and internal procedures to answer to the increase demand of health professionals and citizens in general for better, timely delivered and accurate information (Delpierre et al., 2004; Rahimi and Vimarlund, 2007).

Successful ICT Projects

A comprehensive literature review on large-scale ICT projects executed in several countries/regions identified a set of CSFs (Nagi et al., 2008). Of these, some are recurrently cited in ICT healthcare studies, namely (Hung et al., 2014): (1) Information systems adjustment (Jessup & Valacich, 2008); (2) Business process adjustments (Chatterjee & Segars, 2006); (3) Organizational resistance (Kamhawi, 2007); (4) Top management support (Gil-García & Pardo, 2005; Liu et al., 2010); and (5) The capability of key team members (Siakas & Siakas, 2008; Turner & Lloyd-Walker, 2008; Yang, Kang & Mason, 2008).

Amongst the several PSC that are mentioned, 'iron triangle' seems to be the basis of the most of ICT project success approaches, because fulfilling the criteria for the completion of a project on time, within budget and according to performance specifications are easy to assess (Agarwal & Rathod, 2006; Fortune & White, 2006; Jugdev & Müller, 2005). These classic criteria remain as the most widely used measure of project success and its main value is in offering a simple, direct measure of performance of the project, but it neglects whether the deliverables fulfilled the objectives of the project (Ika, 2009).

According to Sabherwal et al., (2006), ICT projects success in healthcare sector also refers to user satisfaction, system use, perceived usefulness and system quality. PSC have evolved from simple quantifiable 'iron triangle' criteria, which primarily are related to project efficiency (Bryde & Brown, 2005), to measures that have a longer-term perspective directly relating to effectiveness, business success, organizational impact and stakeholder's satisfaction (Atkinson, 1999; Ika, 2009; Jugdev et al., 2001; Müller & Turner, 2007; Shenhar et al., 1997, 2001; Shenhar & Dvir, 2007; Turner & Zolin, 2012).

Success in the strategic use of Healthcare ICT projects relies on the integration of all systems, such as patient records, clinical decision support, transaction processing, digital imaging, and information reporting (Jensen, 2013).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE TRENDS

Today's technology plays a significant role, permitting the storage and rapid retrieval of clinical patient records and other important data/information. At the same time, patients expect that their sensitive personal information to be handled appropriately, to ensure accuracy and confidentiality.

Hospitals are complex organizations and this complexity magnifies the opportunity for inevitable human errors (Weick & Sutcliffe, 2001). A poorly integrated healthcare application system can increase the frequency of medical errors, decrease operational efficiency, and reduce the quality of healthcare services (Themistocleous et al., 2009).

The selected publications emphasized two main aspects regarding the development of ICT initiatives in the healthcare sector; The slowness of adoption of these initiatives and the resistance to change from the part of professionals, mainly physicians (Leonard, 2000, 2004). According to Robinson (2007), appropriate leadership, good communication, detailed roadmap of implementation, setting realistic and measurable goals and specific attention to the active involvement of human resources in terms of motivation and training were considered the most important factors affecting the success of the implementation of these systems.

Extant research largely focuses on the consequences for ICT projects from management support (Dong et al., 2009). Top management support has been argued to be the most critical factor for ICT project success (Liu et al., 2010). Lorenzi & Riley (2003) claimed that organizations need to have three types of skills to produce successful projects, such as: (1) Technical skills - which include a broad range of skills, such as the technical knowledge, experience, and abilities; (2) Project management skills - which include the knowledge, techniques, and skills necessary to manage successfully the ICT projects; and (3) People and organizational skills - which include the wide range of skills necessary to effectively interface with all the ICT stakeholders.

Miller and Oliver (2015) highlighted that the participation of different stakeholders' groups in the design and development process can be essential for a project's success. Reyes-Alcázar et al. (2012) identified several CSFs for the success of ICT projects: (1) Leadership - the importance on improving the quality of healthcare (West et al., 2004); (2) Team work - A multidisciplinary process focused on a healthcare team that shares common goals (Mickan, 2005); (3) Autonomy and responsibility - the need for more autonomy amongst health professionals (Harrison & Dowswell, 2002); (4) An integrated view of healthcare - The quality of patient care as perceived by end-users is a key element (Torres Olivera, 2003); (5) Professional skills - Promoting skills encourages professional development (Reyes-Alcázar et al., 2012); (6) Results focused - The measurement and evaluation of clinical performance, hospital management and end-user satisfaction (Kariami, 2015); (7) Internal and external audits - the concept of continuous improvement of quality (Hyrkäs & Lehti, 2003).

Healthcare processes rely heavily on both data/information and knowledge. In this context, management information plays a crucial role. Numerous studies have shown the positive effects of using ICT in healthcare (Lenz & Reichert, 2007). The support of these systems provides an important asset for specialized services, which increases the quality and safety of patient care by minimizing the likelihood of medical errors (Low & Chen, 2012). Supported on ICT, the complex process of medical decision making could be significantly improved in several ways (Lenz & Reichert, 2007), such as: (1) Contributing to improve data quality, such as timeliness or completeness (Van Walraven, et al., 2004); (2) Contributing to better monitor the current state of a patient (Bates et al., 2001); (3) Detecting mismatches between existing guidelines and the actual patient treatment process (Shiffman, 2004); (4) Generating reminders

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

to ensure that planned actions are not forgotten; (5) Helping to calculate drug dosage from previously entered data (Ambrisko & Nemeth, 2004); (6) Calculating disease probabilities (Hejlesen, 2005).

The benefits that of a good ICT system can bring to healthcare organizations are: (1) Citizens' better access to health systems; (2) Reduce the risk of errors due to lack of data/information; (3) Reduce the time required for the provision of clinical reports; (4) Reduce costs by avoiding repetitive efforts and unnecessary resources.The challenges to be addressed in order to successfully implement these solutions are: (1) Interoperability; (2) Timeliness, integrity, availability and confidentiality of data/information; (3) Alignment of software with sound principles of applications architecture and content management structures; (4) Establishment of rules that allow the user informed consent to the access of their data by professionals; (5) Ensure how to overcome the tremendous difficulties of a change management process.

An important challenge for the future is to seek for a real clinical integration of systems. Clinical integration between providers and hospitals has historically been a goal, still continually sought, but rarely achieved. It will become crucial that the design of future applications be more easily integrated into existing systems, some of them legacy systems, through open communication interfaces that comply with known standards.

CONCLUSION

The investments in ICT implementations have the potential to dramatically change the way individuals, professionals or even society perceive the healthcare sector, and also provide huge opportunities for supporting healthcare professionals, and for improving effectiveness and efficiency in this relevant sector.

Healthcare ICT projects are different from other types of projects in several aspects, such as: the environment in which they are implemented, the diversity of systems and devices they need to work with, the integration and the interoperability between different systems and health organizations, the diversity of interests of the professional groups, the increasing pressure from end-users to an easier access to healthcare services, and the governments pressure to cut costs of subsidizing the public health sector.

Most of times the success or the failure of a ICT project does not only rely on technological factors, but on many other factors, such as, the governance model, the policy environment, multi sectorial involvement of stakeholders and effective human resource management. ICT healthcare projects are generally made up of a significant number of complex tasks and processes, and they depend largely on the quality of existing data/information. A major obstacle is that the diverse players in the health ecosystem traditionally run their own applications with no interconnection among applications. The result is a disintegrated view of how patient care and citizens health information is delivered across the health ecosystem. This situation does not allow obtaining the advantages of a holistic view of a patient's health.

ICT implementation in healthcare need to have well-defined terminologies and standards for interoperability, along with comply with legal requirements. Ideally, these systems should improve quality of care, reduce costs, while maintaining long-standing beneficial patterns of communication, collaboration, quality, safety, and healthcare delivery. ICT projects in healthcare are complex and very difficult to reach a full agreement about what is success. Achieving a successful change will be easier if all stakeholders are appropriately engaged and motivated around the same goal.

In summary, ICT projects in healthcare can lead the industry toward an interconnected healthcare platform by delivering an optimized technology infrastructure combining traditional, cloud-based, and hybrid computing models, pulling together the disparate parts of the healthcare ecosystem, thereby enabling better care, workforce mobility, and security-enhanced data-delivery models.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Abouzhara, M. (2011). Causes of failure in Healthcare IT projects. In *Proceedings of the 3th International Conference on Advance Management Science, IPEDR, 19, IACSIT*, Singapore. Retrieved from http://ipedr.com/vol19/9-ICAMS2011-A00018.pdf

Agarwal, N., & Rathod, U. (2006). Defining success for software projects: An exploratory revelation. *International Journal of Project Management*, *24*(4), 358–370. doi:10.1016/j.ijproman.2005.11.009

Al-Ahmad, A., Al-Fagih, K., Khanfar, K., Alsamara, K., Abuleil, S., & Abu-Salem, H. (2009). A taxonomy of an IT project failure: Root causes. *International Management Review*, *5*(1), 93–104.

Allen, M. J., Kaufmann, D. M., Barret, A., Paterson, D., Sargeant, J., & McLeod, R. (2000). Self-reported effects of computer workshops on physicians' computer use. *The Journal of Continuing Education in the Health Professions*, *20*(1), 20–26. doi:10.1002/chp.1340200105 PMID:11232068

Ambrisko, T., & Nemeth, T. (2004). A computer program for calculating of doses and prices of injectable medications based on body weight or body surface area. *Canadian Journal of Veterinary Research*, *68*, 62–65. PMID:14979437

Ammenwerth, E., Brender, J., Nykänen, P., Prokosch, H. U., Rigby, M., & Talmon, J. (2004). Vision and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. *International Journal of Medical Informatics*, *73*(6), 479–491. doi:10.1016/j.ijmedinf.2004.04.004 PMID:15171977

Ammenwerth, E., & de Keizer, N. (2005). An inventory of evaluation studies of information technology in healthcare: Trends in evaluation research 1982–2002. *Methods of Information in Medicine*, *44*, 44–56. doi:10.1055-0038-1633922 PMID:15778794

Atkinson, R. (1999). Project management: Cost, time and quality, two best guesses and a phenomenon, it's time to accept other success criteria. *International Journal of Project Management*, *17*(6), 337–342. doi:10.1016/S0263-7863(98)00069-6

Bates, D., Cohen, M., Leape, L., Overhage, J., Shabot, M., & Sheridan, T. (2001). Reducing the frequency of errors in medicine using information technology. *Journal of the American Medical Informatics Association*, *8*(4), 299–308. doi:10.1136/jamia.2001.0080299 PMID:11418536

Bates, D., & Gawande, A. (2002). Error in medicine: What have we learned? *Annals of Internal Medicine*, *132*(9), 763–767. doi:10.7326/0003-4819-132-9-200005020-00025 PMID:10787381

Brandrud, A. S., Schreiner, A., Hjortdahl, P., Helljesen, G. S., Nyen, B., & Nelson, E. C. (2011). Three Success Factors for Continual Improvement in Healthcare: An Analysis of the Reports of Improvement Team Members. *BMJ Quality & Safety*, *20*(3), 251–259. doi:10.1136/bmjqs.2009.038604 PMID:21209149

Brender, J., Ammenwerth, E., Nykänen, P., & Talmon, J. (2006). Factors influencing success and failure of health informatics systems: A pilot Delphi study. *Methods of Information in Medicine*, *45*(01), 125–136. doi:10.1055-0038-1634049 PMID:16482383

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Bryde, D., & Brown, D. (2005). The influence of a project performance measurement system on the success of a contract for maintaining motorways and trunk roads. *Project Management Journal*, *35*(4), 57–65. doi:10.1177/875697280403500407

Bukachi, F., & Pakenham-Walsh, N. (2007). Information technology for health in developing countries. *Chest Journal*, *132*(5), 1624–1630. doi:10.1378/chest.07-1760 PMID:17998362

Chatterjee, D., & Segars, A. (2006). Assessing E-Business Capabilities and Effectiveness: A Set of Key E-Business Metrics. *International Journal of Informatics*, *13*(3), 275–295.

Cliff, B. (2012). Using Technology to Enhance Patient-Centered Care. *Journal of Healthcare Management*, *57*(September/October). PMID:23087992

Cooke-Davies, T., & Arzymanow, A. (2003). The maturity of project management in different industries: An investigation into variations between project management models. *International Journal of Project Management*, *21*(6), 471–478. doi:10.1016/S0263-7863(02)00084-4

Crawford, J. (2006). The Project Management Maturity Model. *Information Systems Management*, *23*(4), 50–58. doi:10.1201/1078.10580530/46352.23.4.20060901/95113.7

de Bruin, T., & Rosemann, M. (2005). Towards a Business Process Management Maturity Model. In D Bartmann, F Rajola, J Kallinikos, D Avison, R Winter, & P Ein-Dor, (Eds.), *Proceedings of the 13th European Conference on Information Systems (ECIS 2005)*, Germany, Regensburg.

Delpierre, C., Cuzin, L., Fillaux, J., Alvarez, M., Massip, P., & Lang, T. (2004). A systematic review of computer-based patient record systems and quality of care: More randomized clinical trials or a broader approach? *International Journal for Quality in Health Care*, *16*(5), 407–416. doi:10.1093/intqhc/mzh064 PMID:15375102

Dong, L., Neufeld, D., & Higgins, C. (2009). Top management support of enterprise systems implementations. *Journal of Information Technology*, *24*(1), 55–80. doi:10.1057/jit.2008.21

Drury, P. (2005). The eHealth agenda for developing countries. *World Hospitals and Health Services*, *41*, 38–40. PMID:16512063

Dzenowagis, J. (2005). *Connecting for health: global vision, local insight. Report for the World Summit on the Information Society*. Geneva, Switzerland: WHO Press. Retrieved from http://apps.who.int/iris/bitstream/10665/43385/1/9241593903_eng.pdf

EC. (2009). *Ageing report: Economic and Budgetary projections for the EU-27 Member States (2008-2060).* Luxembourg: Directorate General for Economic and Financial Affairs - DG ECFIN, European Commission, Office for Official Publications of the European Communities. Retrieved from http://ec.europa.eu/economy_finance/publications/pages/publication14992_en.pdf

Elder, L., & Clarke, M. (2007). Past, present and future: Experiences and lessons from telehealth projects. *Open Medicine: a Peer-Reviewed, Independent, Open-Access Journal*, *1*(3), 166–170. PMID:21673948

Eurostat. (2005). *Population projections 2004-2050: EU25 population rises until 2025, then falls*. Retrieved from http://ec.europa.eu/eurostat/documents/2995521/5182986/3-08042005-AP-EN.PDF/ed9d0e20-f75b-4214-b4b2-cf274f7be0d2?version=1.0

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fortune, J., & White, D. (2006). Framing of project success critical success factors by a system model. *International Journal of Project Management*, *24*(1), 53–65. doi:10.1016/j.ijproman.2005.07.004

Gardner, R. M., & Shabot, M. M. (2001). Patient-monitoring Systems. In E.H. Shortliffe, L.E. Perreault, G. Wiederhold & L.M. Fagan (Eds.), Biomedical Informatics (pp. 561-591). New York: Springer Science+Business Media, LLC.

Gauld, R. (2007). Public sector information systems project failures: Lessons from a New Zealand hospital organization. *Government Information Quarterly*, *24*(1), 102–114. doi:10.1016/j.giq.2006.02.010

Gawande, A., & Bates, D. (2000). The use of information technology in improving medical performance. Part II. Physician-support tools. *Medscape General Medicine*, *2*, E13. PMID:11104459

Gheorghiu, F. (2006). Why companies fail on the way to implementing project management methodology. *Project Management Today*, *8*(10), 1–7.

Gil-García, J. R., & Pardo, T. A. (2005). E-government success factors: Mapping practical tools to theoretical foundations. *Government Information Quarterly*, *22*(2), 187–216. doi:10.1016/j.giq.2005.02.001

Gomes, J. (2018). *Organisational Maturity and Information Systems and Technology Projects in Healthcare: The Mediation of Project Management* [PhD dissertation]. Universidade de Lisboa.

Harrison, S., & Dowswell, G. (2002). Autonomy and bureaucratic accountability in primary care: What English general practitioners say. *Sociology of Health & Illness*, *24*(2), 208–226. doi:10.1111/1467-9566.00291

Haux, R. (2006). Health information systems – past, present, future. *International Journal of Medical Informatics*, *75*(3-4), 268–281. doi:10.1016/j.ijmedinf.2005.08.002 PMID:16169771

Häyrinen, K., Saranto, K., & Nykänen, P. (2008). Definition, Structure, Content, Use and Impacts of Electronic Health Records: A Review of the Research Literature. *International Journal of Medical Informatics*, *77*(5), 291–304. doi:10.1016/j.ijmedinf.2007.09.001 PMID:17951106

Heeks, R. (2006). Health Information Systems: Failure, success and improvisation. *International Journal of Medical Informatics*, *75*(2), 125–137. doi:10.1016/j.ijmedinf.2005.07.024 PMID:16112893

Heeks, R. (2008). *Success and failure rates of e-Government in developing/transitional countries: Overview. In Building Digital Opportunities Programme, Information Exchange, University of Manchester's Institute for Development Policy and Management*. UK: Department for International Developments.

Hejlesen, O., Olesen, K., Dessau, R., Beltoft, I., & Trangeled, M. (2005). Decision support for diagnosis of lyme disease. *Studies in Health Technology and Informatics*, *116*, 205–210. PMID:16160260

Hung, S.-Y., Chen, C., & Wang, K.-H. (2014). Critical Success Factors for the Implementation of Integrated Healthcare Information Systems Projects: An Organizational Fit Perspective. *Communications of the Association for Information Systems*, *34*, 775–796. Retrieved from http://aisel.aisnet.org/cais/vol34/iss1/39

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Hyrkäs, K., & Lehti, K. (2003). Continuous quality improvement through team supervision supported by continuous self-monitoring of work and systematic patient feedback. *Journal of Nursing Management*, *11*(3), 208–226. doi:10.1046/j.1365-2834.2003.00369.x PMID:12694365

Ibbs, C., & Kwak, Y. (2000). Assessing project management maturity. *Project Management Journal*, *31*(1), 32–43. doi:10.1177/875697280003100106

Ika, L. A. (2009). Project success as a topic in project management journals. *Project Management Journal*, *40*(4), 6–19. doi:10.1002/pmj.20137

Jachimowicz, V. (2003). Project management maturity model. *Project Management Journal*, *34*(1), 55–56. doi:10.1177/875697280303400108

Jensen, T. B. (2013). Design Principles for Achieving Integrated Healthcare Information Systems. *Health Informatics Journal*, *19*(1), 29–45. doi:10.1177/1460458212448890 PMID:23486824

Jessup, L., & Valacich, J. (2008). *Information Systems Today* (3rd ed.). New Jersey, NJ: Prentice Hall.

Jiang, J., Klein, G., Hwang, H., Huang, J., & Hung, S. (2004). An exploration of the relationship between software development process maturity and project performance. *Information & Management*, *41*(3), 279–288. doi:10.1016/S0378-7206(03)00052-1

Jiang, J., Muhanna, W., & Klenin, G. (2000). User resistance and strategies for promoting acceptance across system types. *Information & Management*, *37*(1), 25–36. doi:10.1016/S0378-7206(99)00032-4

Jorgenson, D., & Stiroh, K. (2000). Raising the speed limit: US economic growth in the information age. In Jorgenson (eds) *Economic growth in the information age*, Cambridge, Massachusetts: MIT Press. Retrieved from https://www.brookings.edu/bpea-articles/raising-the-speed-limit-u-s-economic-growth-in-the-information-age/

Jugdev, K., & Müller, R. (2005). A retrospective look at our evolving understanding of project success. *Project Management Journal*, *36*(4), 19–31. doi:10.1177/875697280503600403

Jugdev, K., Thomas, J., & Delisle, C. (2001). Rethinking project management: Old truths and new insights. *Project Management Journal*, *7*, 36–43.

Kamhawi, E. M. (2007). Critical Factors for Implementation Success of ERP Systems: An Empirical Investigation from Bahrain. *International Journal of Enterprise Information Systems*, *3*(2), 34–48. doi:10.4018/jeis.2007040103

Kaplan, B., & Harris-Salamone, K. (2009). Health IT success and failure: Recommendations from literature and an AMIA workshop. *Journal of the American Medical Informatics Association*, *16*(3), 291–299. doi:10.1197/jamia.M2997 PMID:19261935

Karimi, F., Poo, D., & Tan, Y. (2015). Clinical information systems end user satisfaction: The expectations and needs congruencies effects. *Journal of Biomedical Informatics*, *53*, 342–354. doi:10.1016/j.jbi.2014.12.008 PMID:25542853

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kaushal, R., Shojania, K., & Bates, D. (2003). Effects of computerized physician order entry and clinical decision support systems on medication safety. a systematic review. *Archives of Internal Medicine*, *163*(12), 1409–1416. doi:10.1001/archinte.163.12.1409 PMID:12824090

Kohn, L., Corrigan, J., & Donaldson, M. (Eds.). (2000). *To err is Human: Building a Safer Health System*. Washington: National Academy Press.

Krist, A. H., & Woolf, S. H. (2011). A Vision for Patient-Centered Health Information Systems. *Journal of the American Medical Association*, *305*(3), 300–301. doi:10.1001/jama.2010.2011 PMID:21245186

Lenz, R., & Reichert, M. (2007). IT support for healthcare processes: Premises, challenges, perspectives. *Data & Knowledge Engineering*, *61*(1), 39–58. doi:10.1016/j.datak.2006.04.007

Leonard, K. J. (2000). Information Systems for Healthcare: Why we haven't had more success: The Top 15 Reasons. *Healthcare Management Forum*, *13*(3), 45–51. doi:10.1016/S0840-4704(10)60776-4 PMID:15892319

Leonard, K. J. (2004). Critical Success Factors Relating to Healthcare's Adoption of New Technology: A Guide to Increasing the Likelihood of Successful Implementation. *ElectronicHealthcare*, *2*(4), 72–81.

Lewis, D., Hodge, N., Gamage, D., & Whittaker, M. (2011). Understanding the role of technology in health information systems. *Pacific Health Dialog*, *18*(1), 144–154. PMID:23240349

Liu, S., Zhang, J., Keil, M., & Chen, T. (2010). Comparing senior executive and project manager perceptions of IT project risk: A Chinese Delphi study. *Information Systems Journal*, *20*(4), 319–355. doi:10.1111/j.1365-2575.2009.00333.x

Lorenzi, M., & Riley, R. (2003). Organizational issues = change. *International Journal of Medical Informatics*, *69*(2-3), 97–203. doi:10.1016/S1386-5056(02)00105-3 PMID:12810124

Low, C., & Chen, Y. (2012). Criteria for the evaluation of a cloud- based hospital information system outsourcing provider. *Journal of Medical Systems*, *36*(6), 3543–3553. doi:10.100710916-012-9829-z PMID:22366976

Lucas, H. (2008). Information and communications technology for future health systems in developing countries. *Social Medicine (Social Medicine Publication Group)*, *66*(10), 2122–2132. doi:10.1016/j.socscimed.2008.01.033 PMID:18343005

Mäenpää, T., Suominen, T., Asikainen, P., Maass, M., & Rostila, I. (2009). The outcomes of regional healthcare information systems in health care: A review of the research literature. *International Journal of Medical Informatics*, *78*(11), 757–771. doi:10.1016/j.ijmedinf.2009.07.001 PMID:19656719

Mettler, T., Rohner, P., & Winter, R. (2010). Towards a Classification of Maturity Models in Information Systems. In A. D'Atri (Eds.), *Management of the Interconnected World* (pp. 333–340). doi:10.1007/978-3-7908-2404-9_39

Mickan, S. (2005). Evaluating the effectiveness of health care teams. *Australian Health Review*, *29*(2), 211–217. doi:10.1071/AH050211 PMID:15865572

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Miller, D., & Oliver, M. (2015). *Engaging Stakeholders for Project Success* [White paper]. Project Management Institute. Retrieved from http://www.gbd.dk/files/984_engagingstakeholders.pdf

Müller, R., & Turner, R. (2007). Matching the project manager's leadership style to project type. *International Journal of Project Management*, *25*(1), 21–32. doi:10.1016/j.ijproman.2006.04.003

Munthe, C. (2008). The Goals of Public Health: An Integrated, Multidimensional Model. *Public Health Ethics*, *1*(1), 39–52. doi:10.1093/phe/phn006

Nagi, E., Law, C., & Wat, F. (2008). Examining the Critical Success Factors in the Adoption of Enterprise Resource Planning. *Computers in Industry*, *59*(6), 548–564. doi:10.1016/j.compind.2007.12.001

Paré, G., & Sicotte, G. (2001). Information technology sophistication in health care: An instrument validation study among Canadian hospitals. *International Journal of Medical Informatics*, *63*(3), 205–223. doi:10.1016/S1386-5056(01)00178-2 PMID:11502433

Peabody, J. W., Luck, J., Jain, S., Bertenthal, D., & Glassman, P. (2004). Assessing the Accuracy of Administrative Data in Health Information Systems. *Medical Care, 42*, 1066-1072. Retrieved from http://www.jstor.org/stable/4640857

Rahimi, B., & Vimarlund, V. (2007). Methods to evaluate health information systems in health care settings: A literature review. *Journal of Medical Systems*, *35*(5), 397–432. doi:10.100710916-007-9082-z PMID:17918694

Reyes-Alcázar, V., Torres-Olivera, A., Núñes-García, D., & Almuedo-Paz, A. (2012). Critical Success Factors for Quality Assurance in Healthcare Organisations. In M. Savsar (Ed.), *Quality Assurance Management*. doi:10.5772/33081

Robinson, C. (2007). Clinician adoption of healthcare information technology. *Canadian Nursing Informatics*, *2*(1), 4–21.

Sabherwal, R., Jeyaraj, A., & Chowa, C. (2006). Information System Success: Individual and Organizational Determinants. *Management Science*, *52*(12), 1849–1864. doi:10.1287/mnsc.1060.0583

Schultz, J. (2006). Measuring Service Industry Performance: Some Basic Concepts. *Performance Improvement*, *45*(4), 13–19. doi:10.1002/pfi.2006.4930450405

Sergeant, R., Hatcher, C., Trigunarsyah, B., Coffey, V., & Kraatz, J. (2010). Creating value in Project Management using PRINCE 2. *Queensland University of Technology, Australia and Office of Government Commerce*, UK. Retrieved from https://eprints.qut.edu.au/52853/

Shenhar, A., Dvir, D., Levy, O., & Maltz, A. C. (2001). Project Success: A Multidimensional Strategic Concept. *Long Range Planning*, *34*(6), 699–725. doi:10.1016/S0024-6301(01)00097-8

Shenhar, A., Levy, O., & Dvir, D. (1997). Mapping the dimensions of project success. *Project Management Journal*, *28*, 5–13.

Shenhar, A. J., & Dvir, D. (2007). *Reinventing Project Management: The Diamond Approach to Successful Growth and Innovation*. Boston, MA: Harvard Business School Press.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Shiffman, R., Michel, G., Essaihi, A., & Thornquist, E. (2004). Bridging the guideline implementation gap: A systematic, document-centered approach to guideline implementation. *Journal of the American Medical Informatics Association*, *11*(5), 418–426. doi:10.1197/jamia.M1444 PMID:15187061

Siakas, K. V., & Siakas, E. (2008). The Need for Trust Relationships to Enable Successful Virtual Team Collaboration in Software Outsourcing. *International Journal of Technology Policy and Management*, *8*(1), 59–75. doi:10.1504/IJTPM.2008.016181

Skulmoski, G. (2001). Project maturity and competence interface. *Cost Engineering (Morgantown, W. Va.)*, *43*, 11–18.

Snyder, C. F., Wu, A. W., Miller, R. S., Jensen, R. E., Bantug, E. T., & Wolff, A. C. (2011). The role of informatics in promoting patient-centered care. *Cancer Journal*, *17*(4), 211–218. doi:10.1097/PPO.0b013e318225ff89 PMID:21799327

Sultan, N., & Sultan, Z. (2012). The application of utility ICT in healthcare management and life science research: a new market for a disruptive innovation? In *Proceedings of EURAM 2012 Conference - Social Innovation for Competitiveness, Organizational Performance and Human Excellence*, June 6-8, *Rotterdam School of Management*, Rotterdam.

Themistocleous, M., Mantzana, V., & Morabito, V. (2009). Achieving Knowledge Management Integration through EAI: A Case Study from Healthcare Sector. *International Journal of Technology Management*, *47*(1/2/3), 114–126. doi:10.1504/IJTM.2009.024117

Thomas, J., & Mullaly, M. (2008). *Researching the value of project management*. Philadelphia, PA: Project Management Institute.

Torres Olivera, A. (2003). La gestión por procesos asistenciales integrales: Una estrategia necesaria. *Atencion Primaria*, *31*(9), 561–563. doi:10.1016/S0212-6567(03)79216-6 PMID:12783744

Turner, J. R., & Zolin, R. (2012). Forecasting success on large projects: developing reliable scales to predict multiple perspectives by multiple stakeholders over multiple time frames. *Project Management Journal*, *43*(5), 87–99. doi:10.1002/pmj.21289

Turner, R., & Lloyd-Walker, B. (2008). Emotional Intelligence (EI) Capabilities Training: Can IT Develop EI in Project Teams? *International Journal of Managing Projects in Business*, *1*(4), 512–534. doi:10.1108/17538370810906237

Van der Meijden, M., Tange, H., Troost, J., & Hasman, A. (2003). Determinants of success of inpatient clinical information systems: A literature review. *Journal of the American Medical Informatics Association*, *10*(3), 235–243. doi:10.1197/jamia.M1094 PMID:12626373

Van Walraven, C., Mamdani, M., Fang, J., & Austin, P. (2004). Continuity of care and patient outcomes after hospital discharge. *Journal of General Internal Medicine*, *19*(6), 624–631. doi:10.1111/j.1525-1497.2004.30082.x PMID:15209600

Ward, J., & Daniel, E. (2012). *Benefits Management: How to increase the Business Value of Your IT Projects* (2nd ed.). Chichester, UK: John Wiley and Sons. doi:10.1002/9781119208242

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Weick, K., & Sutcliffe, K. (2001). *Managing the unexpected: Assuring high performance in an age of complexity*. San Francisco, CA: Jossey-Bass.

West, B., Lyon, M., McBain, M., & Gass, J. (2004). Evaluation of a clinical leadership initiative. *Nursing Standard*, *19*(5), 33–41. doi:10.7748/ns.19.5.33.s61 PMID:15524254

Wetering, R., & Batenburg, R. (2009). A PACS maturity model: A systematic meta-analytic review on maturation and evolvability of PACS in hospital enterprise. *International Journal of Medical Informatics*, *78*(2), 127–140. doi:10.1016/j.ijmedinf.2008.06.010 PMID:18752988

WHO. (2006). *Building foundations for eHealth: Progress of Member States.* Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/goe/publications/build_foundations/en/

WHO. (2011). *Global Health and Ageing*. National Institutes of Health publication no. 1. Geneva, Switzerland: National Institute on Aging and National Institutes of Health, World Health Organization. Retrieved from http://www.who.int/ageing/publications/global_health.pdf doi:1.7737/oct.2011

WHO. (2015). *World report on ageing and health.* Geneva, Switzerland: World Health Organization, WHO Press. Retrieved from http://apps.who.int/iris/bitstream/10665/186463/1/9789240694811_eng.pdf?ua=1

Yang, H., Kang, H., & Mason, R. M. (2008). An exploratory study on meta skills in software development teams: antecedent cooperation skills and personality for shared mental models. *European Journal of Information Systems*, *17*(1), 47–61. doi:10.1057/palgrave.ejis.3000730

Yeo, K. T. (2002). Critical failure factors in information system projects. *International Journal of Project Management*, *20*(3), 241–246. doi:10.1016/S0263-7863(01)00075-8

Young, M., Young, R., & Zapata, J. (2014). Project, programme and portfolio maturity: A case study of Australian Federal Government. *International Journal of Managing Projects in Business*, *7*(2), 215–230. doi:10.1108/IJMPB-08-2013-0034

This research was previously published in the International Journal of Privacy and Health Information Management (IJPHIM), 6(2); edited by Francesco Longo and Letizia Nicoletti; ; pages 72-83; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 102
Robotics in Surgical Techniques Robotics in Surgical Techniques: Present and Future Trends

Vassilia Costarides
National Technical University of Athens, Greece

Kostas Giokas
National Technical University of Athens, Greece

Apollon Zygomalas
University of Patras, Greece

Dimitris Koutsouris
National Technical University of Athens, Greece

ABSTRACT

Healthcare robotic applications are a growing trend due to rapid demographic changes that affect healthcare systems, professionals and quality of life indicators, for the elderly, the injured and the disabled. Current technological advances in robotic systems offer an exciting field for medical research, as the interdisciplinary approach of robotics in healthcare and specifically in surgery is continuously gaining ground. This chapter features a review of current applications, from external large scale robotic devices to nanoscale swarm robots programmed to interact on a cellular level.

INTRODUCTION

Since the beginning of human history, surgical operations even in their primitive form have been on the frontier of medicine, due to the lack of non-invasive diagnostic techniques and the traumatic nature of the majority of health issues resulting from centuries of wars and battles. Surgeon, a term introduced in the 14th century means 'a person who heals by manual operation on the patient and has its roots in the Latin world chirurgiae from Greek *cheirourgiki (χειρουργική) and* more specifically their – hand and ergon – work. Surgery, the physical manipulation of the body aimed at the diagnosis, prevention and eventually treatment of diseases and traumas and from early on, the most critical factors that had to be dealt with, were bleeding, pain and infections.

DOI: 10.4018/978-1-7998-1204-3.ch102

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Among the first surgical operations were laceration suturing, amputations, cauterization of wounds and trepanation with evidence such as prehistoric remains in cave paintings and ancient Greek writers' descriptions. In ancient Egypt engravings of surgical tools and papyri with recipes and pharmacopeia for the treatment of numerous diseases have been discovered and in China, physicians performed surgery with the aid of anesthesia, 1600 years before it was adopted in Europe. The evolution and incorporation of anatomy in medical education in the middle ages and Renaissance provided the basis for the systematic research that established the discipline of modern scientific surgery, as a common and safe practice.

Nowadays challenges such as pain management and infections are addressed with the latest technological advances in anesthesia, antisepsis, and novel diagnostic procedures. In our never ending quest for lower mortality rates, less pain and quality services, current economic and social trends guide surgical advances to less invasive techniques that eventually lead to lower hospitalization and recovery times, thus ensuring profit and quality of life to all relevant stakeholders and patients relatively.

BACKGROUND

In the delicate balance between minimally invasive surgery and high-cost state of the art technology, lies robot-assisted surgery, a field rapidly evolving in conjunction with other disciplines, such as micro and nano machines, wireless communications, bioengineering and information technology.

Minimally invasive procedures refer to techniques that target at limiting the incision site, so that wound healing times, pain and infection risk are controlled and minimized. It is widely used by interventional radiologists, employing imaging techniques and catheters, used as interventional instruments directed at the targeted vessel or organ, but also by general surgeons with laparoscopic procedures.

There are several techniques that surgeons use to perform minimally invasive surgery with the patient under general or local anesthesia, depending on the procedure. Usually, laparoscopic instruments either have a telescopic rod lens system that is connected to a camera, or the laparoscope is digital and the charged – couple device is placed at the end of it (Swanstrom, 2013).

The demand of robots in healthcare and specifically in surgery comes from the finite human capability to perform certain actions with adequate efficiency. As described above, latest trends in healthcare align towards automatization, robot – assisted surgery and minimization and therefore efficiency must be maximized along with patient safety, security and sensitivity. The manipulation of robotic instruments via a console, directly or indirectly that translate the surgeon movements to precision actions, was first introduced in 1983 with the Arthrobot (Paul, 1992), a bone mountable hip arthroplasty surgery robot and in 1985 the PUMA 560 (Piltan, 2012) was introduced, for a needle brain biopsy, under CT guidance. In 1988 it was used to perform a transurethral resection and a year earlier for the first laparoscopic cholecystectomy. In 1992 the Robodoc was introduced to mill out precise fittings in the femur for hip replacement (Schulz, 2007).

IDENTIFIED NEEDS

Robot-assisted surgery originally developed to overcome certain limitations of minimally invasive procedures. Based on the analysis of sensor information, healthcare robots use systems capable of doing mechatronic actions to provide healthcare services, such as diagnosis, interventions, treatment, support

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

rehabilitation and prevention (Bogue, 2011). Patients, healthcare professionals and relevant stakeholders such as public institutions, insurance companies, and researchers, are setting the clear lines around the needs and challenges regarding technological evolution and turning points in robot-assisted surgery.

Safety is the most important requirement from patient's point of view and at the same time, professionals like doctors and nurses must maintain safety because of their obligation due to the nature of their profession. Therefore, procedures performed by robotic systems or with their assistance must ensure safety to both patient and healthcare professionals. In general, robots offer newer, better and safer treatments compared to the traditional approaches in many procedures (Finkelstein, 2010).

Even though quality improvements in diagnosis and treatment are crucial and recognized as a need, this must never outweigh safety issues as well as accuracy and consistency in delivering stable, reproducible and predictable results. Surgical accuracy and efficiency, as well as quick recovery times, are important factors, as they are directly affecting indicators of overall efficiency of health systems, and consequently improving the quality of life for patients and working conditions for healthcare professionals.

A specific need that initially triggered the development of telemedicine is medical care in remote areas on earth or even in space missions and even disaster scenarios, where medical facilities as we know them may not be available (Zygomalas, 2015). It is important therefore to develop and evolve flexible and portable, autonomous or remotely controlled semi-autonomous robotic surgical systems that can function in the above cases. From macro-world to micro and macro again, robotic-assisted surgery offers the possibility to access otherwise inaccessible human body areas, without major incisions and perform delicate operations, even on a cellular level in the future.

In the era of big data and information technology, robot-assisted surgical interventions have the possibility to provide detailed information about each case, which can be analyzed and contribute to further research. Furthermore, extended documentation of the procedures provides traceability, improved surgical and recovery time, and consequently lower costs which make health care accessible to all social classes in a long-term perspective.

CLASSIFICATION OF SURGICAL ROBOTS

Several classification systems have been suggested for robotic systems in healthcare in general (Kim, 2016), most distinguishing between imaging, rehabilitation, simulation, prostheses and surgical assistance technologies. The hypothesis behind classification lies on pillars such as potential benefits, technological challenges, and configuration. Factors such as accuracy, space constrictions, compatibility and extension issues, user experience and interface, application specific functionalities, size, sterilization and imaging (CT, C-Arm, MRI) compatibility, sensing and haptic technological challenges, degree of autonomy (passive, active, teleoperative and interactive systems) are also taken into consideration, allowing for future clinical and technological advances to be explored.

Two major categories of surgical robots for general abdominal surgery can be identified, distinguishing in relation to the positioning of the main robotic system and the effective arms with the medical instruments in relation to the human body and the targeted surgical field (Zygomalas, 2015); A) External robotic systems, and B) Internal robotic systems.

The external surgical robots have their basic operating mechanism outside of the surgical field, yet instrumentation is inside the abdomen (or the target area), regardless of the incision approach, which can either be through multiple incisions, one incision (Laparo-Endoscopic Single-Site), or through

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

the body's natural orifices (Natural Orifice Transluminar Endoscopic Surgery) (Rosen, 2011). The external workstations are usually large installations that have to be situated inside the operating room. External endoscopic workstations, specifically oriented towards NOTES surgical techniques have been developed, as well as flexible endoscopic systems (Bardaro, 2006) that allow navigation and surgery in tight workspaces. As it is mentioned above, current trends in surgery are incision minimization with quality improvement and external robotic systems, have been under interdisciplinary research, in terms of miniaturization, in research activities and implementation, offering a new area for development with miniaturized external robots (Sanchez, 2013).

Regarding internal robotic systems, miniaturized laparoscopic systems, that are located entirely inside the human body, have been described and developed (Lehman, 2007; Oleynikov, 2005) that are introduced into the abdominal cavity via various methods and once positioned, function completely inside the abdomen. For applications in NOTES techniques, internal endoscopic miniature systems, that can be swallowed and self- assembled have been proposed (Assembling Reconfigurable Endoluminal Surgical system, 2014).

In addition to mini-robotic devices, endoscopic swallowable capsules have been developed that allow for imaging, telemetry, and even histology if provided with appropriate biopsy tools (Saurin, 2016). Finally, the evolution of nanotechnology contributes positively to surgical technologies, with the introduction of swarm intelligence robotics, interacting on a cellular and intra-cellular level (Mavroidis, 2013).

External Laparoscopic Robotic Systems

Laparoscopic surgery in which operations are performed through small incisions (usually 0.5–1.5 cm) reduces patient trauma and diminishes recovery times but at the same time limits surgeon's dexterity and vision because of the degrees of freedom that are lost due to the entry incisions, the rigidity of the laparoscopic instruments and the two-dimensional visual feedback.

The best known robotic system today is the da Vinci® Surgical System (Intuitive Surgical (2002) which incorporates a magnified vision system with a 3D high-definition view of the area and robotic technology that allows surgeons' hands movements to be translated into smaller, precise movements of its tiny wristed instruments inside the patient's body. Currently, there are several surgical systems under development that are marketed as daVinci's rivals, such as the SportTM Surgical System (Titanic Medical, 2016), the Telelap AlF-X (TransEnterix, 2016), the Raven Surgical robot (Kehoe, 2014) and the Araknes project (Sanchez, 2013) among others.

The daVinci surgical system is a proprietary technology, consisting of the surgeon's console, a patient-side cart, a vision system and instruments. The surgeon operates from the console, viewing the 3D image of the surgical field. It seamlessly translates the surgeon's movements in real time into precise, movements of the surgical instruments. More specifically, the 3D image featured in the console, provides HD vision and natural colors, up to 10x magnification, and the patented instruments have 7 degrees of freedom and a range of motion, greater than the human hand. The surgeon's hand tremors are minimized, and the console features built-in microphone in order to facilitate operating room communication.

The laparoscopic robot arm is equipped with a 3D camera and a light at the end. The camera sends the images from the abdominal cavity to the video monitor in the operation room and the surgeon console. The surgeon is at all times 100% in control of the system. Since its operation it has brought minimally invasive surgery to more than 3 million patients, covering various medical specialties such as cardiac, colorectal, general, gynecologic, head & neck, thoracic, and urologic surgery.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The SPORT™ Surgical System, which is yet under development comprises a robotic platform, surgeon – control for a single incision that incorporates a 3D HD vision system and multi articulating instruments for minimally invasive surgery. There is also a surgeon workstation that provides surgeons with an interface for controlling the instruments as well as the 3D view during the minimally invasive procedures. It is a surgical system, designed to provide the multi-articulating instruments and the 3D vision system for insertion into the patient's body cavity through a single incision of approximately 25mm with the surgeon being able to control it from the workstation. The instruments are designed to be coupled with sterile single use tools so that quality is ensured. The robotic platform is also designed to include a mast, a boom and wheels for optimal configurability for a variety of surgical indications and the ability to be maneuvered around the operating room and surgical centers where applicable.

The Telelap ALF-X surgical system, acquired by Transenterix Inc., is granted a CE mark, intended to be used for laparoscopic surgery in the abdomen and pelvis (Stark, 2015). It features haptic feedback, which leads to improved safety and allows for certain maneuvers to be performed with greater confidence. The system also tracks the surgeon's eye movements, positioning the camera, so the field of view is centered where the eyes are looking. The system is highly adjustable for an optimal ergonomic experience.

The SurgiBot System, by TransEnterix as well is designed to be a minimally invasive surgical platform that allows multiple instruments to be introduced and deployed through a single incision. The flexible instruments may be controlled directly by the surgeon from inside the sterile field, with robotic assistance. Its arm stabilizes the camera and alleviates the need for a dedicated person to hold it, and the instruments were developed to minimize incisions.

The Raven Surgical Robotic System was developed by the universities of Berkeley, Davis, and Santa Cruz mainly. It is a state-of-the-art open architecture (software & hardware) surgical robotic system and a prototype system including four arms which is currently available at the Bionics Lab (UCSC), while several other identical systems are to be distributed among leading groups across the US (Harvard, Hopkins, U. Nebraska, UCLA, UCB, UCSC, UW). The system uses a new statistical approach, Apprenticeship learning in terms of robot learning, targeted at allowing at robotic surgical assistants a higher level of autonomy, by being able to execute specific subtasks with performance exceeding human potential in regards to speed and smoothness. The latest version of the Raven Surgical System is IV, which was developed at the University of California – Santa Cruz and it includes two cameras and four robotic arms, and it facilitates a collaborative effort of two surgeons interacting with the surgical site in teleportation. The surgical arms are based on a spherical mechanism with a remote center located at the point of entry of the tool into the human body. Two surgeons in different remote locations can be connected via UDP protocol, as it is designed in the system architecture, with the use of a proprietary client interface and thus remotely operate the robots in the master/slave standard configuration. In experiments, several fundamental laparoscopic surgeries were completed while using a sampling rate of 1kHz.

Miniature External Robotic Systems

The Araknes Project, under the Large-scale integrating project (IP), FP7-ICT-Challenge 3: Components, systems and engineering/Micro/nanosystems funding scheme (Sanchez, 2013), is targeted at engaging micro and nanotechnologies from on-bench applications to in vivo surgical scenarios in endoluminal surgical approaches, thus further reducing operative trauma and enhancing therapeutic outcome of minimally invasive procedures. The system is based on an array of micro robotic instrumentation with components that include a flexible oro-pharyngo-esophageal access port and an umbilical access port,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

assistive and operative micro-robots allowing a bi-manual operation, an imaging system consisting of an endoscopic camera for global vision and a stereoscopic camera for visualizing the operative area, biosensors for tissue analysis (normally not available during laparoscopy) and an operating console incorporating the image of the operative field.

The developed robot under the project - ARAKNES SPRINT, Single-Port Laparoscopy bimanual robot - is yet a single incision master-slave robotic platform in terms of configuration, which consists of two high dexterity anthropomorphic manipulators, a stereoscopic camera and a master console for the robot control by the surgeon.

Even though external surgical robots have been an established reality in the medical community, there are still facing many challenges, mostly of technological and financial nature that must be addressed in the years to come, with an interdisciplinary approach and methodology (Vitiello, 2013; Marcus, 2013). In order for such an expensive technology to be incorporated in the services of health care systems, extensive cost- benefit analyses and needs assessments must be performed in mentioned environments, so that potential benefits can be substantiated.

At all times, the pre-operative diagnostic data must be matched in the surgical field so that localization of pathology is accurate. Furthermore, the tissue deformation that occurs by the manipulation of the tissue with either tools or physiological processes must be modeled and characterized in real time so that pathology may be identified in real time and deformations predicted. Haptic environments and tools may offer significant benefit together with visual technologies that will improve the restricted vision provided from current cameras.

Above all, as surgical robots are designed for all case scenarios, there is the size factor, that dictates new systems to be light-weight (versus current bulky systems) and with flexible access platforms (Noonan, 2011) that incorporate actuators and sensors.

Internal Laparoscopic Robotic Systems

Internal surgical laparoscopic robots aim at overcoming the obstacles of the classic and even the robotic laparoscopic surgery by including systems that augment vision and improve dexterity, using externally controlling robots that are completely located inside a body cavity. In vivo miniature robots, with either fixed-base or mobile platform, that fit entirely inside the peritoneal cavity, are now a novel technology offering a different approach (Rentschen, 2004; Zygomalas, 2015).

When it comes to fixed based laparoscopic robots, the main purpose is to provide an enhanced view of the surgical field, by augmenting vision and depth perception. Robots are placed with classical laparoscopic tools and are possible to be moved throughout the procedure, without the need to change the sterile field and make a new incision. Miniature PT (pan tilt) cameras have been developed as initial work allowing for panning of 360o and tilting of 45o, rotating two independent axes and with the use of LED lights (Oleynikov, 2004). Such robots were inserted within the standard laparoscopic port into the peritoneal cavity and had provided additional views of the cavity.

Laparoscopic robots, which feature a movable platform, remotely controlled, consist of systems that provide forward, reverse and turning motion (Rentschen, 2006). There is usually a mechanism that allows for counter rotation prevention and wheels are incorporated so that it is possible to maneuver on organs, like liver, spleen and bowels without causing damage. Similar robotic mobile laparoscopic systems have been developed (Rentschen, 2006) that feature adjustable focus cameras to improve depth perception within the peritoneal cavity, thus eliminating the need for camera port incisions eventually.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Biopsy instrumentation in addition to vision assistance has been added in several systems for one-port laparoscopic biopsy. Forceps providing adequate biopsy samples in porcine subjects have demonstrated promising results, with the onboard camera used to locate the biopsy site as the robot traverses the peritoneal cavity (Rentschen, 2007).

Apart from fixed base camera and wheeled robots, magnetic drive robots for biopsy, vision, and task assistance has been developed (Lehman, 2007), and tested in vivo. These robots assisted surgeons in laparoscopic surgery by providing manipulators with a dexterity not limited by the entry incision, alongside with an enhanced view.

Internal Endoscopic Robotic Systems

In advancing minimally invasive surgery an evolving field that initiated interest in the medical community since 2005 is Natural Orifice Transluminal Endoscopic Surgery. NOTES features operations where scarless abdominal operations are performed with the use of endoscopic systems or miniaturized robots that are passed through the natural orifices of the human body and then through internal incisions, therefore avoiding external incisions and scars. NOTES is the integration of laparoscopic surgical techniques with endoscopic instrumentations and the current technological focus of the development of flexible or autonomous wireless platforms that incorporate robotic technology, thus using micro or nanoscale robotic devices providing capabilities such as imaging, retraction, tissue and organ manipulation, and precise maneuverability in the abdominal cavity. NOTES can be classified according to primary routes of access as transvaginal, transrectal and trans gastric (Kaloo, 2004; Marescaux, 2007).

The transvaginal entrance has already been established as a beneficial technique, as it is performed under direct visualization and through 2010 in 64% of the cases, rigid endoscopes were used in hybrid procedures (Auyang, 2011). Transrectal procedures are less common (Bulian, 2014) and trona gastric access has been less frequently used, as it had the technical challenge of gaining peritoneal access and a challenging gastrostomy closure.

Swallowable microsystems, which can be reassembled in the stomach cavity are proposed and developed, with potential to change their configuration according to location and task assigned, while monitored by navigation systems and controlled via wireless communication consoles operated by surgeons or healthcare professionals (Nagy, 2009). Within the ARES project (Assembling Reconfigurable Endoluminal Surgical system, 2014) such a system was developed, that can be disassembled after the operation, into individual models or reconfigured for a consequent intervention, eventually reassembled into a snake type structure. The modular nature of the proposed system can incorporate various extra modules, even during operations or dispose of unnecessary modules that can be discarded or replaced, with the use of magnets as the driving mechanism of self-assembly. Magnets can provide a reversible connection – disconnection platform due to the nature of the phenomenon per se and can enhance reconfiguration options. Modular mini robots for possible use in NOTES or laparoendoscopic single site (LESS) surgery, have also been investigated in-silico using professional robotics simulation software with promising results (Zygomalas, 2011, 2012, 2014).

Magnetic anchoring and guidance systems (MAGS), that usually consist of an external magnet and an intracorporeal device with a camera or surgical instrumentation, have been proposed to be used for NOTES. The idea features insertion of the internal part through either an incision or a natural orifice and coupling via magnetic fields, across the skin tissue layer. The use of cameras with LED-wrapped trocars (Park, 2007) and the insertion of cautery and retractors modules have been tested in experimental models (Zeltser, 2007).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Endoluminar miniaturized robots are a much promising technological field for minimally invasive surgery, even though for the time being, implementation in real practice is limited mainly due to size constraints (Zygomalas, 2015).

ROBOTIC CAPSULES

In addition to miniaturized robotic devices that are endoscopically inserted into the abdominal cavity, capsule endoscopy, is a trend that has been an uprising in current medical research and it has started from imaging and telemetry capsules, used for diagnostic issues. A robotic system for capsule endoscopy has modules for power, vision, telemetry, localization and diagnosis and/or manipulation. Robotic capsules are now incorporating locomotion capabilities and are evolving into devices able to perform interventional tasks. One of the most critical issues is the development of an adequate propulsion system that will be delicate enough not to damage tissues, yet powerful enough to advance in complex tissue structures and perform surgical tasks with accuracy, safety and efficiency. The propulsion system is part of the overall design of the capsule that has a critical role in the successful deployment of the device. It can either be internal, external or hybrid where more than one propulsion mechanisms are used. Various methods have been proposed according to the geometry of the target area, like leg or paddle like mechanisms, activated by micro motors, magnetic actuators, or wall-to-wall motion with the aid of micro-propellers.

When it comes to internal propulsion mechanisms, the robotic system is more autonomous, having better control of its mobility. Legged propulsion systems have already been developed (Gorini, 2006), extending from 3 to 12 legged as nature is providing inspiration with bio-mimetic (bio – inspired) systems, that can vary from earthworm/inchworm-like that are based on piezo actuators or shape memory alloy springs (Kim, 2004, 2005, 2009). In such nature-inspired propulsion systems, the modules are to be swallowed and then reassembled inside the gastrointestinal track, using magnetic or wireless communication systems (Kim, 2013; Fang, 2015).

Cilia and flagella, cell organelles, similar in structure but with different functions, provide different motions, with the motion of flagella undulating and wave-like, whereas cilia motion is more complicated, 3D and with a power and recovery stroke. Both have been suggested and used as the basis for propulsion mechanisms of robotic endovascular systems (Li, 2006; Feng, 2007; Kosa, 2007) while other systems have been created with paddling based propulsion robots, that mimic canoeist paddling (Park, 2006, 2007). These feature linear actuators with dual cylinders that perform a cyclic motion, which helps the robot to advance in cycles.

Apart from external mechanisms, it is possible to use the electrical stimuli produced by the natural contractions of the intestinal smooth muscles that can be applied via electrodes to the robot (Woo, 2009, 2011). The peristalsis that occurs (the natural movement of the bowels) creates a force that allows potentially for a robot to move opposite to the contraction end of the lumen and it can be bidirectional depending on which electrode is activated.

Vibratory propulsion is a concept where the robot has an eccentric rotating motor that generates a (Carta, 2011; Zabulis, 2008) centripetal force thus allowing for the horizontal component of the force to propel the robot.

In the principle of internal reaction propulsion, the robot moves by the reaction force caused by the movement of an internal mass, without legs or wheels as in other systems (Liu, 2008), exploiting the piezoelectric phenomenon, incorporating a linearly-moving mass generate the reaction force. The most

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

common mechanism for external propulsion, which bypasses the issue of internal actuators, releasing space for other instrumentation, is external propulsion by means of an external magnetic field interacting with internal magnetic components. It can be either guided by an MRI device, with its magnetic field interacting with potential coils incorporated in the robot or by an external permanent magnet that can be operated by a healthcare professional or even a robotic navigation system (Lien, 2012).

In order to combine advantages of internal and external propulsion systems and reduce risks and disadvantages mentioned above, hybrid systems have been developed that combine magnetic and motor mechanisms (Menciassi, 2009) with motion achieved by external magnetic fields and orientation by internal mechanisms. Other hybrid systems, bring together legged actuators with external magnetic dragging (Simi, 2010).

With the microelectromechanical systems advances, it is possible to integrate into miniature robotic capsules several tools like sensors for pH measurements, pressure and temperature sensors and blood detection platforms. Integration of interventional tools, is still under development. However interventions such as biopsy, suturing, clipping and drug delivery, have already been tested in vitro or in vivo in porcine models (Iddan, 2000).

NANOTECHNOLOGY AND SURGICAL ROBOTS

Since the miniaturization of electromechanical systems (MEMS) and their incorporation in general surgery (Rebelo, 2004), the next goal set by science was to advance further on the nanoscale, thus manipulation of the matter on an atomic, molecular, and supramolecular scale, with at least one dimension sized from 1 to 100 nanometers. Nanotechnology in medical applications and more specifically in surgery has been part of research, from drug delivery systems to diagnostics and interventions. Nanorobots, controllable electromechanical systems, are capable of actuation, sensing, communication, propulsion, signal and information processing and interaction with tissue. They can also manifest swam behavior and have a unique potential in delivering clinical results of high quality and accuracy (Mavroidis, 2013).

Apart from diagnosis and treatment of diseases with nanoscale delivery systems (Lenaghan, 2013; Dadkhah, 2013), monitoring and treatment of chronic diseases (Cavalcanti, 2008) or even incisionless surgery is the next potential application. In order for the nanorobots to be functional, several conditions must be met. Initially, the swarm intelligence (Bonabeau, 1999), which is the collective behavior of nanorobots that can self-organize them, should be accompanied by appropriate propulsion so that movement and navigation inside the human body can be realized. For sensors and actuators, communications, control, and decision-making functions, it is important to incorporate a power supply system, capable of supporting the described functions for the estimated time of the surgical intervention.

Nanorobots will be scaled to fit inside a drop of saline and then can be programmed to move to specific locations, to areas of interest and interact with cells directly, instead of manipulating tissue in a macro scale. This interaction on a cellular level means that cells can be guided to either die or regenerate with simple commands, inspired by computer programming. The input taken from cells will serve as logical input that can train and enhance logical function and cognitive behavior of the nanorobots, sustaining and advancing their autonomous intelligence.

Apart from collecting information and interacting with living cells, nanorobots can be integrated or even compiled of compatible biomaterials and serve as scaffolds in surgical interventions that require bone or soft tissue reconstruction.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

As the world marches towards a new era of intelligent technologies, interacting autonomously within an environment of shared information and big data, it is but certain that self-aware cognitive robotic systems will be incorporated into everyday medical practice. Minimally invasive surgeries, performed with micro and nano-robots, will be ensuring minimal tissue damage, high accuracy, and quality, as technological advances will be incorporated into the discipline, along with accumulated experience. New capabilities and procedures will be developed, and it is likely that human horizon and possibilities will expand including thought controlled devices and what once was science fiction, will become mundane reality.

REFERENCES

Assembling Reconfigurable Endoluminal Surgical System. (2014). CORDIS. Retrieved June 8, 2016, from http://cordis.europa.eu/publication/rcn/10899_en.html

Auyang, E. D., Santos, B. F., Enter, D. H., Hungness, E. S., & Soper, N. J. (2011). Natural orifice translumenal endoscopic surgery (NOTES®): A technical review. *Surgical Endoscopy*, *25*(10), 3135–3148. doi:10.100700464-011-1718-x PMID:21553172

Bardaro, S. J., & Swanström, L. (2006). Development of advanced endoscopes for natural orifice transluminal endoscopic surgery (NOTES). *Minimally Invasive Therapy & Allied Technologies*, *15*(6), 378–383. doi:10.1080/13645700601038069 PMID:17190663

Bogue, R. (2011). Robots in healthcare. *Industrial Robot: An International Journal*, *38*(3), 218–223. doi:10.1108/01439911111122699

Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm intelligence: from natural to artificial systems (No. 1)*. Oxford University Press.

Bulian, D. R., Runkel, N., Burghardt, J., Lamade, W., Butters, M., Utech, M., ... Lehmann, K. S. (2014). Natural Orifice Transluminal Endoscopic Surgery (NOTES) for colon resections—analysis of the first 139 patients of the German NOTES Registry (GNR). *International Journal of Colorectal Disease*, *29*(7), 853–861. doi:10.100700384-014-1883-1 PMID:24798629

Carta, R., Sfakiotakis, M., Pateromichelakis, N., Thoné, J., Tsakiris, D. P., & Puers, R. (2011). A multi-coil inductive powering system for an endoscopic capsule with vibratory actuation. *Sensors and Actuators. A, Physical*, *172*(1), 253–258. doi:10.1016/j.sna.2011.03.036

Cavalcanti, A., Shirinzadeh, B., & Kretly, L. C. (2008). Medical nanorobotics for diabetes control. *Nanomedicine; Nanotechnology, Biology, and Medicine*, *4*(2), 127–138. doi:10.1016/j.nano.2008.03.001 PMID:18455965

Dadkhah, M., Kumar, N., & Yoon, J. (2013). Design and simulation of a 3D actuation system for magnetic nano-particles delivery system. In *Intelligent Robotics and Applications* (pp. 177–187). Springer Berlin Heidelberg. doi:10.1007/978-3-642-40852-6_20

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Fang, H., Wang, C., Li, S., Wang, K. W., & Xu, J. (2015). A comprehensive study on the locomotion characteristics of a metameric earthworm-like robot. *Multibody System Dynamics*, *35*(2), 153–177. doi:10.100711044-014-9428-5

Feng, Y., Li, W., Li, M., & Sun, L. (2007, December). Structure optimization of the endoscopic robot ciliary leg based on dimensional analysis. In *Robotics and Biomimetics, 2007. ROBIO 2007. IEEE International Conference on* (pp. 109-114). IEEE.

Finkelstein, J., Eckersberger, E., Sadri, H., Taneja, S. S., Lepor, H., & Djavan, B. (2010). Open Versus Laparoscopic Versus Robot-Assisted Laparoscopic Prostatectomy: The European and US Experience. *Reviews in Urology*, *12*(1), 35–43. PMID:20428292

Gorini, S., Quirini, M., Menciassi, A., Pernorio, G., Stefanini, C., & Dario, P. (2006, February). A novel SMA-based actuator for a legged endoscopic capsule. In *Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on* (pp. 443-449). IEEE. 10.1109/BIOROB.2006.1639128

Iddan, G., Meron, G., Glukhovsky, A., & Swain, P. (2000). Wireless capsule endoscopy. *Nature*, *405*(6785), 417. doi:10.1038/35013140 PMID:10839527

Intuitive Surgical. (2002). *The da Vinci Surgical System*. Retrieved June 8, 2016, from http://www.intuitivesurgical.com/

Kalloo, A. N., Singh, V. K., Jagannath, S. B., Niiyama, H., Hill, S. L., Vaughn, C. A., ... Kantsevoy, S. V. (2004). Flexible transgastric peritoneoscopy: A novel approach to diagnostic and therapeutic interventions in the peritoneal cavity. *Gastrointestinal Endoscopy*, *60*(1), 114–117. doi:10.1016/S0016-5107(04)01309-4 PMID:15229442

Kehoe, B., Kahn, G., Mahler, J., Kim, J. H., Lee, A., Nakagawa, K., . . . Goldberg, K. (2014, May). Autonomous multilateral debridement with the raven surgical robot. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on* (pp. 1432-1439). IEEE. 10.1109/ICRA.2014.6907040

Kim, B., Lee, S., Park, J. H., & Park, J. O. (2004, August). Inchworm-like microrobot for capsule endoscope. In *Robotics and Biomimetics, 2004. ROBIO 2004. IEEE International Conference on* (pp. 458-463). IEEE.

Kim, B., Lee, S., Park, J. H., & Park, J. O. (2005). Design and fabrication of a locomotive mechanism for capsule-type endoscopes using shape memory alloys (SMAs). *Mechatronics, IEEE/ASME Transactions on, 10*(1), 77-86.

Kim, B., Park, S., Jee, C. Y., & Yoon, S. J. (2005, August). An earthworm-like locomotive mechanism for capsule endoscopes. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on* (pp. 2997-3002). IEEE.

Kim, B., Park, S., & Park, J. O. (2009, July). Microrobots for a capsule endoscope. In *Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference on* (pp. 729-734). IEEE. 10.1109/AIM.2009.5229926

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kim, J., Gu, G. M., & Heo, P. (2016). Robotics for Healthcare. In Biomedical Engineering: Frontier Research and Converging Technologies (pp. 489-509). Springer International Publishing. doi:10.1007/978-3-319-21813-7_21

Kim, L., Tang, S. C., & Yoo, S. S. (2013, October). Prototype modular capsule robots for capsule endoscopies. In *Control, Automation and Systems (ICCAS), 2013 13th International Conference on* (pp. 350-354). IEEE. 10.1109/ICCAS.2013.6703922

Kósa, G., Jakab, P., Hata, N., Jólesz, F., Neubach, Z., Shoham, M., . . . Székely, G. (2008, October). Flagellar swimming for medical micro robots: theory, experiments and application. In *Biomedical Robotics and Biomechatronics, 2008. BioRob 2008. 2nd IEEE RAS & EMBS International Conference on* (pp. 258-263). IEEE. 10.1109/BIOROB.2008.4762812

Kósa, G., Shoham, M., & Zaaroor, M. (2007). Propulsion method for swimming microrobots. *Robotics*. *IEEE Transactions on*, *23*(1), 137–150.

Lehman, A. C., Rentschler, M. E., Farritor, S. M., & Oleynikov, D. (2007). The current state of miniature in vivo laparoscopic robotics. *Journal of Robotic Surgery*, *1*(1), 45–49. doi:10.100711701-007-0019-9 PMID:25484938

Lenaghan, S. C., Wang, Y., Xi, N., Fukuda, T., Tarn, T., Hamel, W. R., & Zhang, M. (2013). Grand challenges in bioengineered nanorobotics for cancer therapy. *Biomedical Engineering*. *IEEE Transactions on*, *60*(3), 667–673.

Li, W., Guo, W., Li, M., & Zhu, Y. (2006, June). A novel locomotion principle for endoscopic robot. In *Mechatronics and Automation, Proceedings of the 2006 IEEE International Conference on* (pp. 1658-1662). IEEE. 10.1109/ICMA.2006.257445

Lien, G. S., Liu, C. W., Jiang, J. A., Chuang, C. L., & Teng, M. T. (2012). Magnetic control system targeted for capsule endoscopic operations in the stomach—design, fabrication, and in vitro and ex vivo evaluations. *Biomedical Engineering*. *IEEE Transactions*, *59*(7), 2068–2079.

Liu, Y., Yu, H., & Yang, T. C. (2008). Analysis and Control of a Capsubot. *City*, *10*, 2.

Marcus, H., Nandi, D., Darzi, A., & Yang, G. Z. (2013). Surgical robotics through a keyhole: From today's translational barriers to tomorrow's "disappearing" robots. *Biomedical Engineering*. *IEEE Transactions on*, *60*(3), 674–681.

Marescaux, J., Dallemagne, B., Perretta, S., Wattiez, A., Mutter, D., & Coumaros, D. (2007). Surgery without scars: Report of transluminal cholecystectomy in a human being. *Archives of Surgery (Chicago, Ill.)*, *142*(9), 823–826. doi:10.1001/archsurg.142.9.823 PMID:17875836

Mavroidis, C., & Ferreira, A. (Eds.). (2013). *Nanorobotics: current approaches and techniques*. Springer Science & Business Media. doi:10.1007/978-1-4614-2119-1

Menciassi, A., Valdastri, P., Quaglia, C., Buselli, E., & Dario, P. (2009, September). Wireless steering mechanism with magnetic actuation for an endoscopic capsule. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE* (pp. 1204-1207). IEEE. 10.1109/IEMBS.2009.5332426

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Nagy, Z., Fluckiger, M., Oung, R., Kaliakatsos, I. K., Hawkes, E. W., Nelson, B. J., ... Abbott, J. J. (2009). Assembling reconfigurable endoluminal surgical systems: Opportunities and challenges. *International Journal of Biomechatronics and Biomedical Robotics*, *1*(1), 3–16. doi:10.1504/IJBBR.2009.030054

Noonan, D. P., Vitiello, V., Shang, J., Payne, C. J., & Yang, G. Z. (2011, September). A modular, mechatronic joint design for a flexible access platform for MIS. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on* (pp. 949-954). IEEE. 10.1109/IROS.2011.6094907

Oleynikov, D., Rentschler, M., Hadzialic, A., Dumpert, J., Platt, S. R., & Farritor, S. (2005). Miniature robots can assist in laparoscopic cholecystectomy. *Surgical Endoscopy And Other Interventional Techniques*, *19*(4), 473–476. doi:10.100700464-004-8918-6 PMID:15742124

Park, H., Park, S., Yoon, E., Kim, B., Park, J., & Park, S. (2007, April). Paddling based microrobot for capsule endoscopes. In *Robotics and Automation, 2007 IEEE International Conference on* (pp. 3377-3382). IEEE. 10.1109/ROBOT.2007.363994

Park, S., Bergs, R. A., Eberhart, R., Baker, L., Fernandez, R., & Cadeddu, J. A. (2007). Trocar-less instrumentation for laparoscopy: Magnetic positioning of intra-abdominal camera and retractor. *Annals of Surgery*, *245*(3), 379–384. doi:10.1097/01.sla.0000232518.01447.c7 PMID:17435544

Park, S., Park, H., Park, S., & Kim, B. (2006). A paddling based locomotive mechanism for capsule endoscopes. *Journal of Mechanical Science and Technology, 20*(7), 1012-1018.

Paul, H. A., Bargar, W. L., Mittlestadt, B., Musits, B., Taylor, R. H., Kazanzides, P., ... Hosainpour, A. (2012). PUMA-560 Robot Manipulator Position Computed Torque Control Methods Using MATLAB/SIMULINK and Their Integration into Graduate Nonlinear Control and MATLAB Courses. *International Journal of Robotics and Automation*, *3*(3), 167–191.

Rebello, K. J. (2004). Applications of MEMS in surgery. *Proceedings of the IEEE*, *92*(1), 43–55. doi:10.1109/JPROC.2003.820536

Rentschler, M., Hadzialic, A., & Dumpert, J. (2004). In vivo robots for laparoscopic surgery. *Studies in Health Technology and Informatics*, *98*, 316–322. PMID:15544298

Rentschler, M. E., Dumpert, J., Platt, S. R., Ahmed, S. I., Farritor, S. M., & Oleynikov, D. (2006). Mobile in vivo camera robots provide sole visual feedback for abdominal exploration and cholecystectomy. *Surgical Endoscopy And Other Interventional Techniques*, *20*(1), 135–138. doi:10.100700464-005-0205-7 PMID:16333551

Rentschler, M. E., Dumpert, J., Platt, S. R., Iagnemma, K., Oleynikov, D., & Farritor, S. M. (2006). Modeling, analysis, and experimental study of in vivo wheeled robotic mobility. Mechanical Engineering Faculty Publications.

Rentschler, M. E., Dumpert, J., Platt, S. R., Iagnemma, K., Oleynikov, D., & Farritor, S. M. (2007). An in vivo mobile robot for surgical vision and task assistance. *Journal of Medical Devices*, *1*(1), 23–29. doi:10.1115/1.2355686

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Sanchez, L. A., Le, M. Q., Rabenorosoa, K., Liu, C., Zemiti, N., Poignet, P., & Dario, P. (2013). A case study of safety in the design of surgical robots: The ARAKNES platform. In *Intelligent Autonomous Systems 12* (pp. 121–130). Springer Berlin Heidelberg. doi:10.1007/978-3-642-33932-5_12

Saurin, J. C., Beneche, N., Chambon, C., & Pioche, M. (2016). Challenges and Future of Wireless Capsule Endoscopy. *Clinical Endoscopy, 49*(1), 26.

Schulz, A. P., Seide, K., Queitsch, C., von Haugwitz, A., Meiners, J., Kienast, B., & Jürgens, C. (2007). Results of total hip replacement using the Robodoc surgical assistant system: Clinical outcome and evaluation of complications for 97 procedures. *International Journal of Medical Robotics and Computer Assisted Surgery*, *3*(4), 301–306. doi:10.1002/rcs.161 PMID:18000945

Simi, M., Valdastri, P., Quaglia, C., Menciassi, A., & Dario, P. (2010). Design, fabrication, and testing of a capsule with hybrid locomotion for gastrointestinal tract exploration. *Mechatronics*. *IEEE/ASME Transactions on Mechatronics*, *15*(2), 170–180. doi:10.1109/TMECH.2010.2041244

Stark, M., Pomati, S., D'Ambrosio, A., Giraudi, F., & Gidaro, S. (2015). A new telesurgical platform–preliminary clinical results. *Minimally Invasive Therapy & Allied Technologies*, *24*(1), 31–36. doi:10.3109/13645706.2014.1003945 PMID:25627435

Swanstrom, L. L., & Soper, N. J. (2013). *Mastery of endoscopic and laparoscopic surgery*. Lippincott Williams & Wilkins.

Titanic Medical. (2016). *The SPORT Surgical System*. Retrieved June 8, 2016, from http://www.titanmedicalinc.com/product/

Transenterix. (2016). *The Telelap ALF-X Surgical Robotic System*. Retrieved June 8, 2016, from http://www.alf-x.com/en/

Vitiello, V., Lee, S. L., Cundy, T. P., & Yang, G. Z. (2013). Emerging robotic platforms for minimally invasive surgery. *Biomedical Engineering*. *IEEE Reviews in*, *6*, 111–126.

Woo, S. H., Kim, T. W., Lee, J. H., Kim, P. U., Won, C. H., & Cho, J. H. (2009). Implemented edge shape of an electrical stimulus capsule. *International Journal of Medical Robotics and Computer Assisted Surgery*, *5*(1), 59–65. doi:10.1002/rcs.235 PMID:19170130

Woo, S. H., Kim, T. W., Mohy-Ud-Din, Z., Park, I. Y., & Cho, J. H. (2011). Small intestinal model for electrically propelled capsule endoscopy. *Biomedical Engineering Online*, *10*(1), 108. doi:10.1186/1475-925X-10-108 PMID:22177218

Zabulis, X., Sfakiotakis, M., & Tsakiris, D. P. (2008, August). Effects of vibratory actuation on endoscopic capsule vision. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* (pp. 5901-5904). IEEE. 10.1109/IEMBS.2008.4650557

Zeltser, I. S., Bergs, R., Fernandez, R., Baker, L., Eberhart, R., & Cadeddu, J. A. (2007). Single trocar laparoscopic nephrectomy using magnetic anchoring and guidance system in the porcine model. *The Journal of Urology*, *178*(1), 288–291. doi:10.1016/j.juro.2007.03.001 PMID:17499803

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Zygomalas, A., Giokas, K., & Koutsouris, D. (2012). Modular assembly micro-robots for natural orifice transluminal endoscopic surgery, the future of minimal invasive surgery. *International Journal of Reliable and Quality E-Healthcare*, *1*(4), 43–55. doi:10.4018/ijrqeh.2012100104

Zygomalas, A., Giokas, K., & Koutsouris, D. (2014). In Silico Investigation of a Surgical Interface for Remote Control of Modular Miniature Robots in Minimally Invasive Surgery. *Minimally Invasive Surgery*, *2014*, 307641. doi:10.1155/2014/307641 PMID:25295187

Zygomalas, A., Gkiokas, K., & Koutsouris, D. D. (2011). In silico development and simulation of a modular reconfigurable assembly micro-robot for use in natural orifice transluminal endoscopic surgery. *Hellenic Journal of Surgery*, *83*(4), 190–196. doi:10.100713126-011-0037-y

Zygomalas, A., Kehagias, I., Giokas, K., & Koutsouris, D. (2015). Miniature surgical robots in the era of NOTES and LESS: Dream or reality? *Surgical Innovation*, *22*(1), 97–107. doi:10.1177/1553350614532549 PMID:24828382

KEY TERMS AND DEFINITIONS

Laparoscopy: A type of surgical procedure, in which the incisions are small and through which instruments and a viewing tube are inserted.

Micro – Electro – Mechanical Systems: Mechanical and electromechanical structures and devices that are made with microfabrication techniques and with dimensions that vary from below one micron to several millimeters.

Modular Microrobots: Robotic Systems that consist of diverse several subunits that could be assembled in order to construct a functional miniature robot.

Natural Orifice Transluminal Endoscopic Surgery: A novel surgical technique that uses the natural orifices of the human body in order to approach the peritoneal cavity.

Robotic Surgery: Surgical procedures that incorporate as aid, the latest technological developments in robotic systems.

This research was previously published in Design, Development, and Integration of Reliable Electronic Healthcare Platforms edited by Anastasius Moumtzoglou; pages 0-100; copyright year 2017 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 103
A Next Generation of IT:
A Health Record Example with Terminology Binding

Anthony J. Mallia
Edmond Scientific Company, USA

ABSTRACT

A potential new generation computing environment is emerging which combines wiki technology with semantic web concepts. This has brought about the fusion of the wiki execution ecosystem, a semantic web for model-driven applications, and a high-level language as an extension to wiki text for accelerated development. Semantic MediaWiki provides this platform and a fragment of a health record, including allergy intolerance as structured in HL7 FHIR with terminology bindings to SNOMED CT and to HL7 terminologies was developed by the author in a short timeframe (approximately 10 hours). The system navigates around the health record and controls the entry of terms in the record from controlled ValueSets. All terminologies and ValueSets are integrated into the prototype.

INTRODUCTION

There is a pending information system development revolution. The arrival of the sixth generation of computing is imminent and it is found in the mechanism for developing applications based on two concepts:

1. Open data platforms providing execution of a high-level language (annotation markup) and supporting distributed ecosystems based on wiki technology
2. Emergence of Semantic Web as a knowledge structure for use in ordinary applications (not just artificial intelligence)

The fourth generation of computing has fourth generation language (4GL) support and is aimed at improving development productivity. Enterprise resource planning (ERP) systems typically use 4GL to define system behavior.

DOI: 10.4018/978-1-7998-1204-3.ch103

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The fifth generation is aimed at the field of artificial intelligence based on Semantic Web which introduced the triple store, as well as the management of graph knowledge structures beyond the relational database.

So, in the migration to this next generation we need to be aware of the changes to software construction needed to take advantage of new capabilities and not just apply old programming practices. Specifically, this necessitates the implementation of concepts such as structured data capture in the language of a Semantic Web platform and not in Java (which is not a 4GL).

Part of the motivation for using a Semantic Web platform is the speed with which implementation can be achieved over traditional programming mechanisms. Some precursor development tools exhibited productivity gains in the order of 20 times over traditional development methods Gartner Inc. (2003).

The change of thinking, which must occur between existing programming techniques and those supported by Semantic Web platforms, is demonstrated by how an Electronic Health Record (EHR) could be migrated to this environment.

The paper explores implementation of an EHR in a sixth generation as represented by Semantic MediaWiki. It examines a specific case of the binding of a health care record entry to a term expressing the entry's type as part of a standard terminology where both the record and the terminology are loaded in a Semantic Web environment.

THE CHANGE IN THE BINDING MECHANISM

Existing approaches have binding from a health record entry to a terminology as a multi-string match (a coded element structure including the code system and the term code) to a terminology concept often through a Common Terminology Service (CTS).

The CTS is a service to access the terminology callable by the health record system with legacy programming techniques. One function is that it allows access to the display value of a concept.

The binding example which is described here illustrates the lower capability of existing traditional programming environments compared to the Semantic Web. Data is stored in primitive types – strings, numbers, dates and all relationships are built in the constructed software. Examples of this are found through multiple generations of Health Level 7 [HL7] including the latest standard Fast Healthcare Interoperability Resource [FHIR]. (HL7 FHIR, 2018)

A healthcare record is a complex graph usually anchored with a patient resource. Healthcare actions planned and taken for this patient are recorded in clinical and administrative records (subgraphs) linked to the patient. Identification, coding from terminology systems, and display representation precisely state the records' contents. Examples of terminology systems include Logical Observation Identifiers Names and Codes (LOINC), International Statistical Classification of Diseases and Related Health Problems (ICD), and Systematized Nomenclature of Medicine Clinical Terms (SNOMED International).

The HL7 coded element, as part of the health record entry, supplies a structure containing the unique identity of the coding system (namespace), version of the coding system, unique code, and display text of the semantic. These components will be called system, code, and display. To achieve a precise semantic identification, the system and code components are required. The display is derived from their combination.

Here is a fragment of an HL7 FHIR Allergy Intolerance (HL7 FHIR, 2017) in XML where the entry display is "No Known Drug Allergy."

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Code Fragment 1. (FHIR Allergy Intolerance Code in XML)

```
<AllergyIntolerance xmlns="http://hl7.org/fhir">

  <code>
    <coding>
      <system value="http://snomed.info/sct"/>
      <code value="409137002"/>
      <display value="No Known Drug Allergy"/>
    </coding>
    <text value="NKDA"/>
  </code>
```

The binding is achieved by matching <system> and <code> against the terminology service to extract the display value= “No Known Drug Allergy” which is inserted in the record. This binding must be explicitly implemented in traditional programming.

As expression of terminology moves to Semantic Web and ontologies, the existing binding is not optimal for a semantic web platform - the approach does not make use of link closure between a reference and its target.

While the terminology developers and in particular SNOMED have advocated a health concept be represented by an OWL Class and there has been discussion that an entry in the health record could be an instance of that concept, there has been no discovered literature discussing the specific implementation which would then behave well within a semantic web platform.

If the semantic web platform is potentially the next generation of computing, then it would be important to establish how the health record and the terminologies behave in this environment. This also provides a bridge into clinical decision support where clinical knowledge is also linked to the terminologies and patient record although this topic is not covered in this paper.

In the search of Semantic Web literature references, there are few mentions of its use for Health Records or the binding to terminology except in the HL7 RDF Working group. The expression of terminology in Semantic Web is well referenced but the Health Record side is not. (Rector, Qamar, & Marley, 2009)

No exploration of binding through a proof of concept has been done to the author’s knowledge. The feasibility of this approach is explored in this paper by implementation of a fragment with an available semantic web platform.

WHAT SOLUTIONS EXIST IN LITERATURE

SNOMED CT shows a representation in OWL where the Concept is an OWL Class with a single identifier concatenating the namespace with the term id {sctid}. The OWL representation of SNOMED CT makes use of URIs for identifying Concepts with a structure:

http://snomed.info/id/sctid
(International Health Terminology Standards Developmemt Organization (IHTSDO), 2014)

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Thus, the binding from the health record entry to the concept should have a reference to this single identifier. The prevailing thinking is that the health record entry is an instance of the term as a class which defines its type.

Additional work has been referenced for LexWiki and ICD10 Collaborative Revision Platform (Solbrig, 2008). This is a platform for editing terminologies. In this approach each term has a single name which identifies it. No binding to the Health Record is described although it is assumed that it would be an instance to class relationship. Discussion with the author has confirmed this.

Thus, an instance of Allergy Intolerance associated with a patient can have a link to a type http://snomed.info/id/409137002 (no known drug allergies). The prefix identifies that it is a SNOMED term and the number is the term identifier.

The HealthCare RDF Research

In 2014, the Health Level 7 (HL7) standards development organization formed the "RDF for Semantic Interoperability workgroup" jointly with the World Wide Web Consortium (W3C) to establish semantic interoperability of structured healthcare information. It used the W3C Resource Description Framework (RDF) and related standards to express machine-processable meaning. It supported the HL7 mission to increase the effectiveness and efficiency of healthcare delivery.

The first work in the workgroup was to create an RDF representation of Fast Healthcare Interoperability Resources (HL7 FHIR, 2018). During this activity some prototype FHIR resources in RDF were constructed and imported into Protégé by the author to demonstrate what an RDF representation of a health record might look like on a Semantic Web platform. The current representation of FHIR, as a set of connected resources with their classes, is useful for managing healthcare information (not just exchanging it), and led to the exploration of a representation of the FHIR EHR as a Semantic Web ontology.

Semantic Web mechanisms can be used for low order problems which are not about discovering new knowledge but are about capturing and sharing existing knowledge. The concepts Class, Individual, and Property become the model element types of the description. Triple stores become the database. Thus, a complex graph can be represented by a Semantic Web platform to manage the model, instances, and user navigation in an integrated ecosystem.

In Semantic Web, an instance (individual) can have multiple rdf:type relationships and a Semantic Web platform understands the Class behavior including its sub-classing (rdfs:subClassOf) mechanism used heavily in the SNOMED CT representation.

WHICH KIND OF EVALUATION WAS PERFORMED ON THE SOLUTION

The evaluation of the approach was done by practical demonstration of health record and terminology fragments not only to illustrate the binding mechanism between the health record entry and the terminology which was technically easy, but also to evaluate whether the health record functions could be implemented in the platform without the use of legacy languages.

A mock health record fragment, aligned with HL7 FHIR was created and a subset of SNOMED CT was entered so that the binding would happen automatically by closure in the platform. Display values of the Concepts were available from the terminology to be shown in the health record and allowable concept ValueSets defined by FHIR were presented for selection during structured data capture.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Three use cases were tested:

1. Ability to pick from a ValueSet of permitted terms during structured data capture
2. Display of the human readable term display value from the terminology when viewing the record
3. Ability to display the definition of a term or ValueSet from the terminology when viewing the health record

Decision for the Prototype Implementation

Two implementation options for the prototype were considered

- Traditional programming environment - OWL API (OWL API) and Java
- Semantic MediaWiki (SMW) and Wikitext

Traditional Programming Environment: OWL API and Java

A prototype application involving the definition of a graph ontology and the creation, editing, and reporting of information structured with this ontology was first developed by the author in Protégé using the OWL API.

The user interface would have to be built with Java tools. This is not only a much larger project but did not show the Semantic Web Platform only feasibility.

Semantic MediaWiki (SMW)

A number of applications have been prototyped by the author with SMW using Wikitext, the Wiki markup language to develop a working system.

Although this is not a pure Semantic Web platform with regard to RDF/OWL, it does have export mechanisms to RDF. The platform is very fast when implementing an ontology. It also has reasonable user interface development capabilities for viewing, navigating and structured data capture.

Decision

The example health record was implemented with Wikitext extensions for Semantic MediaWiki (https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki). This early stage proof of concept uses SMW to deploy the ontology both the type definitions and the instances.

This decision is partly influenced by the difference between Protégé and OWL being Open World Assumption based and Semantic MediaWiki being more oriented to a Closed World Assumption where the intention is to constrain allowable values and which is more in line with the governance of an EHR.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

DESCRIPTION OF THE HEALTH RECORD AND TERMINOLOGY FRAGMENTS

Semantic Web Description of FHIR

The HL7 RDF work group (HL7 FHIR/RDF, 2017) chose RDF to be the same structural representation as the other HL7 FHIR formats – XML and JSON so that round trip conversion could be performed.

This RDF example of Code Fragment 1 (FHIR Allergy Intolerance Code) represented in RDF turtle looks like:

Individuals are anonymous (blank nodes) and are nested within some structure (not shown). Thus the matching process was the same as the legacy mechanism with a string match of Coding.system (http://snomed.info/sct) and Coding.code (409137002) against the terminology. At this stage it has not progressed to using the closure mechanism we are expecting.

Semantic Web Description of SNOMED CT

SNOMED International is the trading name of the International Health Terminology Standards Development Organization (IHTSDO). IHTSDO's core product is SNOMED CT, however there are many other supporting products and services that are required in order to ensure the successful adoption of the SNOMED CT which are provided by IHTSDO as secondary services.

SNOMED International released a version of SNOMED CT in Web Ontology Language (OWL). Here is a code fragment showing the SNOMED concept used above:

The sample concept name http://snomed.info/id/409137002, is a concatenation of the system prefix http://snomed.info/id/ and the code 409137002 as defined by SNOMED. The element is an OWL class rather than an instance or individual. The display is an rdfs:label, which is used for human viewing. The concept also uses the assertion in a subclass of http://snomed.info/id/138875005, which is the top-level concept in SNOMED CT as shown:

Code Fragment 2. (RDF Allergy Intolerance Code)

```
[a fhir:AllergyIntolerance;

  fhir: AllergyIntolerance.code [
    fhir:CodeableConcept.coding [
       fhir:Coding.system [ fhir:value "http://snomed.info/sct" ];
       fhir:Coding.code [ fhir:value "409137002" ];
       fhir:Coding.display [ fhir:value " No Known Drug Allergy (situation)" ]
    ]
  ]
];
```

Code Fragment 3. (SNOMED Concept in RDF)

```
<http://snomed.info/id/409137002> rdf:type owl:Class ;
   rdfs:label "No Known Drug Allergies" ;
   rdfs:subClassOf <http://snomed.info/id/138875005> .
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Code Fragment 4. (SNOMED Top Concept)

```
<http://snomed.info/id/138875005> rdf:type owl:Class
    rdfs:label "SNOMED CT Concept"@en .
```

(Author note: in a later version of SNOMED CT there are intermediate classes between the 409137002 term and the top-level concept).

Thus, the SNOMED Ontology is a Class graph with every member being a subclass (not necessarily directly) of the SNOMED CT Concept.

A number of approaches have converged on the single identifier for a concept but there appears to be no existing work to examine what has to happen to the health record to bind to this single identifier.

The binding is thus a single string match which would have to be coded in a traditional programming environment. However, in a semantic web platform this binding is computed by the platform itself so that a graph is known where the Allergy Intolerance instance has closure to a Terminology Class and can be traversed with Semantic Web instructions.

The method for doing this binding was identified in the HL7/W3C RDF working group to be the application of the rdf:type as a semantic class of the CodeableConcept.coding individual (an anonymous individual or blank node) in the record. This then supported both the string match and the closure to the type.

Here is the structure of a record fragment in turtle with the added type “a http://snomed.info/id/409137002”. This is the same as “rdf:type http://snomed.info/id/409137002”). Turtle introduces a special abbreviation for the rdf:type property using the”a”keyword.

Closure will only occur if the full qualified identities match.

There are two options to get the rdf:type if the Coding.system and Coding.code are known:

- Using a reasoner
- Explicit transformation

Code Fragment 5. (Allergy Intolerance code RDF binding to SNOMED CT)

```
[a fhir:AllergyIntolerance;

  fhir: AllergyIntolerance.code [
    fhir:CodeableConcept.coding [a http://snomed.info/id/409137002;
       fhir:Coding.system [ fhir:value "http://snomed.info/sct" ];
       fhir:Coding.code [ fhir:value "409137002" ];
       fhir:Coding.display [ fhir:value "No Known Drug Allergy (situation)" ]
    ]
  ]
]
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Using a Reasoner

Given some additional information, a reasoner can infer the type and cause the closure to happen. This additional information could be provided in a bridging ontology if we do not want to change the record or terminology.

The bridging ontology can declare a general Class axiom which shows the mapping to the Concept Type (or OWL Class as we have been working in OWL) to the values in the coding:

This general class axiom infers that all individuals which have of the restricted property code and system values (defines a class) will be a subclass of the SNOMED class http://snomed.info/id/409137002.

This approach was tested using Protégé and was found to work. It effectively generates the triples needed to navigate the graph and allow every instance of No Known Drug Allergy to be linked to that SNOMED Terminology Class.

However, the volume of statements needed to define the binding is excessive since there needs to be this restriction statement for every member of the terminology – reaching in the hundreds of thousands.

Explicit Transformation

Another way to do the binding is to programmatically insert the rdf:type in the record as needed. This mechanism is preferred given the following discussion of the long-term approach to semantic representation. The other overriding reason is that SMW does not have a reasoner for scalability concerns.

In the proof of concept, the records were entered with the direct binding as shown below. Since we are only dealing with a Semantic Web environment the string matching Codeable Concept can be omitted since we can enter the rdf:type directly.

As will be shown later the Codeable Concept values can be added to the term class.

Code Fragment 6. (Bridging Ontology)

```
[ rdf:type owl:Restriction ;
  rdfs:subClassOf <http://snomed.info/id/409137002> ;
  owl:onProperty fhir:CodeableConcept.coding ;
  owl:someValuesFrom [ rdf:type owl:Class ;
                       owl:intersectionOf ( [ rdf:type owl:Restriction ;
                                              owl:onProperty fhir:Coding.code ;
                                              owl:allValuesFrom [ rdf:type owl:Restriction ;
                                                                  owl:onProperty fhir:value ;
                                                                  owl:hasValue "409137002"
                                                                ]
                                            ]
                                            [ rdf:type owl:Restriction ;
                                              owl:onProperty fhir:Coding.system ;
                                              owl:hasValue <http://snomed.info/sct>
                                            ]
                                          )
                     ]
] .
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Code Fragment 7. (Snomed Binding to the individual)

```
[a fhir:AllergyIntolerance; a http://snomed.info/id/409137002;

   other properties
];
```

The reason the rdf:type can be applied to the whole AllergyIntolerance and not to the object of AllergyIntolerance.code is that the HL7 FHIR meaning of "code" is a classification of the Allergy Intolerance and maps to an rdf:type. This is not the case with all properties of an instance where the semantics of the term need to be declared such as Clinical Status and there will be a specific property for AllergyIntolerance.clinical status having a blank node value that is purely its term as an rdf:type.

ValueSet Representation

ValueSets are constraints on the allowable codings to be used in the context and in HL7 FHIR are aligned to string matching. A value set specifies a set of codes drawn from one or more code systems (HL7 Vocabulary, 2018).

In Semantic Web, the allowable set of codes is a Class based on its set theory origin. The subclasses of the particular ValueSet Class represent the allowable Class identities which can be used in the context of each element of the patient record.

String Matching Declaration

In the Allergy Intolerance example, FHIR defines the ValueSet as:

http://hl7.org/fhir/ValueSet/allergyintolerance-code

This ValueSet (HL7, 2018) includes codes from the following code systems:

Import all the codes that are contained in http://hl7.org/fhir/ValueSet/substance-code
Include codes from http://snomed.info/sct where concept is-a 418038007 (Propensity to adverse reactions to substance)
Include codes from http://snomed.info/sct where concept is-a 267425008 (Lactose intolerance)
Include codes from http://snomed.info/sct where concept is-a 29736007 (Syndrome of carbohydrate intolerance)
Include codes from http://snomed.info/sct where concept is-a 340519003 (Lysine intolerance)
Include codes from http://snomed.info/sct where concept is-a 190753003 (Sucrose intolerance)
Include codes from http://snomed.info/sct where concept is-a 413427002 (Acquired fructose intolerance)
Include codes from http://snomed.info/sct where concept is-a 716186003 (No known allergy)

However, the actual list of codes to be matched is large and must be deployed to achieve the match.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Semantic Web Declaration of ValueSet

All the actual values do not need to be articulated since the term classes may be subclasses of the ones defined in the ValueSet.

The equivalent definition in Semantic Web is:

http://hl7.org/fhir/ValueSet/substance-code
or http://snomed.info/id/418038007 (Propensity to adverse reactions to substance)
or http://snomed.info/id/29736007 (Syndrome of carbohydrate intolerance)
or http://snomed.info/id/340519003 (Lysine intolerance)
or http://snomed.info/id/190753003 (Sucrose intolerance)
or http://snomed.info/id/413427002 (Acquired fructose intolerance)
or http://snomed.info/id/716186003 (No known allergy)

The Allergy Intolerance instance must be a member of the union of the sets defined by the classes. By the previous definition, this union is the class representing the ValueSet.

Combination of Health Record, Terminology and ValueSet

The combination of these ontologies through closure allows semantic navigation from the record through the hierarchies (classes, subclasses) of the terminology ontology. Thus, an adverse reaction to a concept which is a higher concept than a medication type can be detected in the intersection of the class sets of concepts.

Integration of the health record with the terminology allows the practical implementation of ValueSets for restriction of concepts to be applied to the health record during data capture.

However, when we are dealing with a Semantic Web platform we may not have access to the traditional programming mechanisms for identifying and selecting the class URI under the restriction of the ValueSet identified for that element.

The binding from the Record entry is therefore restricted to a subclass of the ValueSet Class.

The approach was tested initially with Protégé which could show the closure only. Then in a fuller user implementation in Semantic MediaWiki, the ValueSets for data capture could be demonstrated with a selection window without resorting to traditional programming techniques.

SEMANTIC MEDIAWIKI OVERVIEW

The introduction of Semantic Web concepts to wiki was achieved by the Semantic MediaWiki (SMW) extension to MediaWiki. In general, historical applications added semantic tagging to wiki pages. However, the tool extension “Page Forms” allows for the collection and manipulation of classical fielded information. Coming from the system development area, the capabilities are equivalent to 4GL. (Kamel Boulos, 2009)

SMW uses the wiki features strong in the areas of user presentation and its high-level annotation markup - Wikitext, was extended by SMW’s open source project based out of Karlsruhe Institute of Technology to support Semantic Web concepts.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

In SMW, the rdf.type is called Category when it refers to the set that the individual belongs to. When rdf:type refers to the metaclass SMW uses Namespace to identify it.

The SMW Wikitext has a powerful in line query function aligned to SPARQL.

Code Fragment 8 is an example of an in-line query to retrieve instances of Category Parks. It displays them using their geographic coordinates in Google Maps:

Code Fragment 9 is an example in a data capture Form showing the selection of a Category from a tree structure of Categories underneath the Top_Category.

The tree structure is displayed from the Category subclass relationships under the Top_Category. This will be shown later in the ValuesSet prototype structured data capture example.

The Semantic MediaWiki Term Binding

In SMW the sample SNOMED CT concept was implemented as:

Category SCT:409137002 (No Known Drug Allergies)

The name has concatenated the System as a prefix "SCT:", the code "409137002", and the display in parentheses. The prefix SCT: represents the URI prefix http://snomed.info/id/.

This was purely pragmatic to get the best use of the platform.

The Category page may contain information as shown in Code Fragment 10 (SNOMED CT Concept in SMW). The [[Category:]] statement defines this category to be a subcategory of SCT: 138875005. The page was designed with fields in a template called Terminology. This provides the breakout of properties System, Code, Display, and Definition which are used in the run time implementation and provide the translation with the traditional Code structure – System and Code.

Code Fragment 8. (SMW #ask Query)

```
{{#ask:[[Category:Parks]] | ?Has coordinates | format=googlemaps}}
```

Code Fragment 9. (Semantic tree entry)

```
{{{field|Category | input type=tree | top category=Top_Category | height=200}}}
```

Code Fragment 10. (SNOMED CT Concept in SMW)

```
{{Terminology
|System=http://snomed.info/sct
|Code=409137002
|Display=No Known Drug Allergies
|Definition=No known history of drug allergy
}}
[[Category:SCT:138875005 (SNOMED CT Concept)]]
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

The Prototype Electronic Health Records (EHR) fragment

The prototype EHR fragment, defined using Unified Modeling Language (UML) Class diagrams, shows links between the instances and the types identified by categories. See Figure 1. UML representation of the ontology.

The graph structure is defined in the class diagram through class attributes and associations and also through subclassing relationships. The example FHIR:AllergyIntolerance declares the structural type along with semantic restrictions. In Semantic Web, class restrictions determine the allowed properties (from their domain). In SMW, the restrictions are applied in a template and form definitions.

A FHIR fragment was translated into a semantic web ontology described in UML before entry to SMW. The FHIR structure ontology is transcribed using category, property, templates and form page SMW editors alongside the health records entered with Forms. Thus, structural types such as Allergy-Intolerance, the SNOMED terminology, and the patient record are all computationally accessible in the SMW environment.

The SMW template declares fields in the page along with the property linking the page to the field. Thus, a "Start Date" field may have a property "Has Start Date" linking the page (instance) of the activity to its start date. The triple store and the value in the page field link a named individual to a simple data type or to a reference to another page. Another page name becomes an HTML link in the user interface.

The editing of the value in the page field causes the triplestore to be edited and the user is not aware of this process although the results can be viewed in the property browser.

The instance identity structure is designer chosen. In this example it has the server location prefix, which must be unique within that server (wiki) see linked data later in the paper. The server prefix is followed by any context qualifier which makes the remainder of the identity unique.

The ontology is expressed in SMW as categories (classes), subcategories, properties (attributes and associations), and instances (sometimes referred to as "articles" in SMW). The structure is documented using UML with explanation. A dependency relationship (dotted line with arrow) is drawn by linking the object (individual) to its classes.

The ontology diagram omits most attributes (data properties) for clarity.

Properties in semantic web are unidirectional in their declaration. However, they may be traversed in both directions with an inverse property. The directional arrow shows the declaration direction. For example, instances of allergy intolerance point to instances of patient using the property FHIR:AllergyIntolerance patient (not the other way). The in-line queries support navigation from the patient to the allergies.

In SMW an inverse property is declared with a negative sign prefix to the property. -FHIR:AllergyIntolerance patient navigates from the Patient to the AllergyIntolerance instance.

FHIR ONTOLOGY PROOF OF CONCEPT IMPLEMENTATION

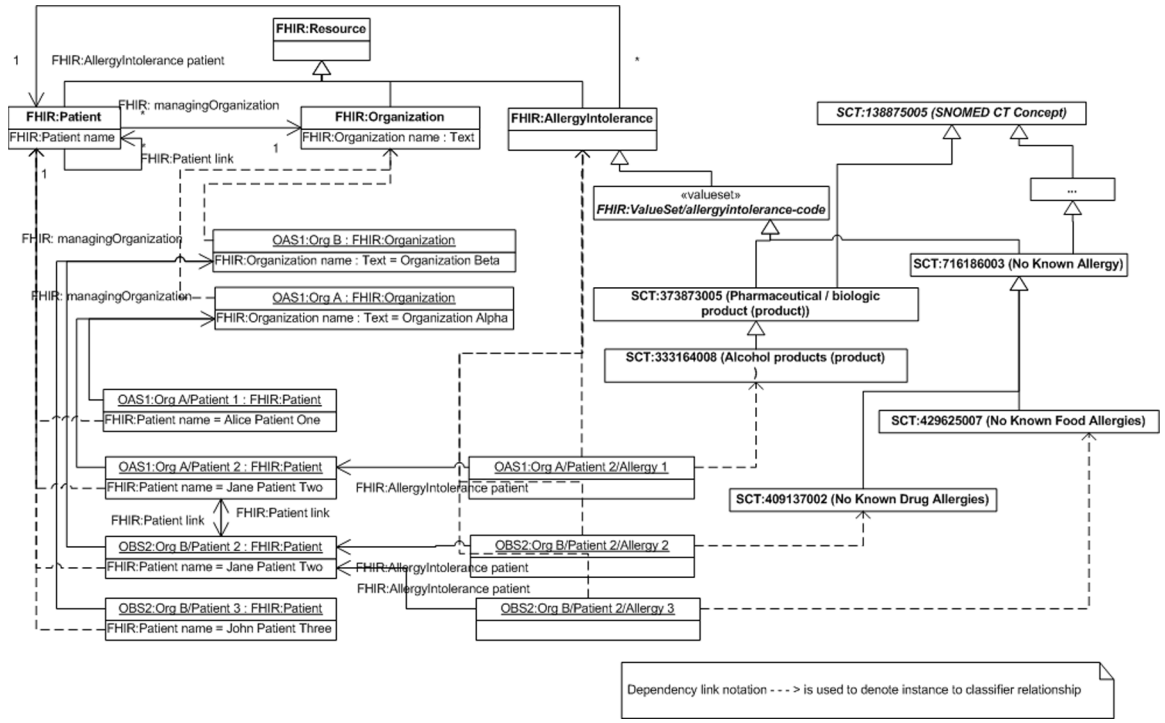
All figures in this section are captured screenshot fragments from SMW (except for the UML diagrams).

The General Ontology

Figure 2 shows the introductory user interface where the contents are generated with in-line queries and includes several patients at different managing organizations. The two managing organizations are

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. UML representation of the ontology



organizations alpha and beta. The records are located on server 1 and server 2. The servers are managed by each organization. They are identified with the prefixes OAS1 and OBS2, which are used as prefixes to define the location of a resource (e.g., OAS1:xxx).

Patients are identified by one or more organizations (assigning authority) where the assigning authority prepends the patient id. OAS1:Org A/Patient 1 has an assigning authority of organization alpha. The record is located at server OAS1.

The properties, reformatted from the HL7 FHIR properties are established to link this together:

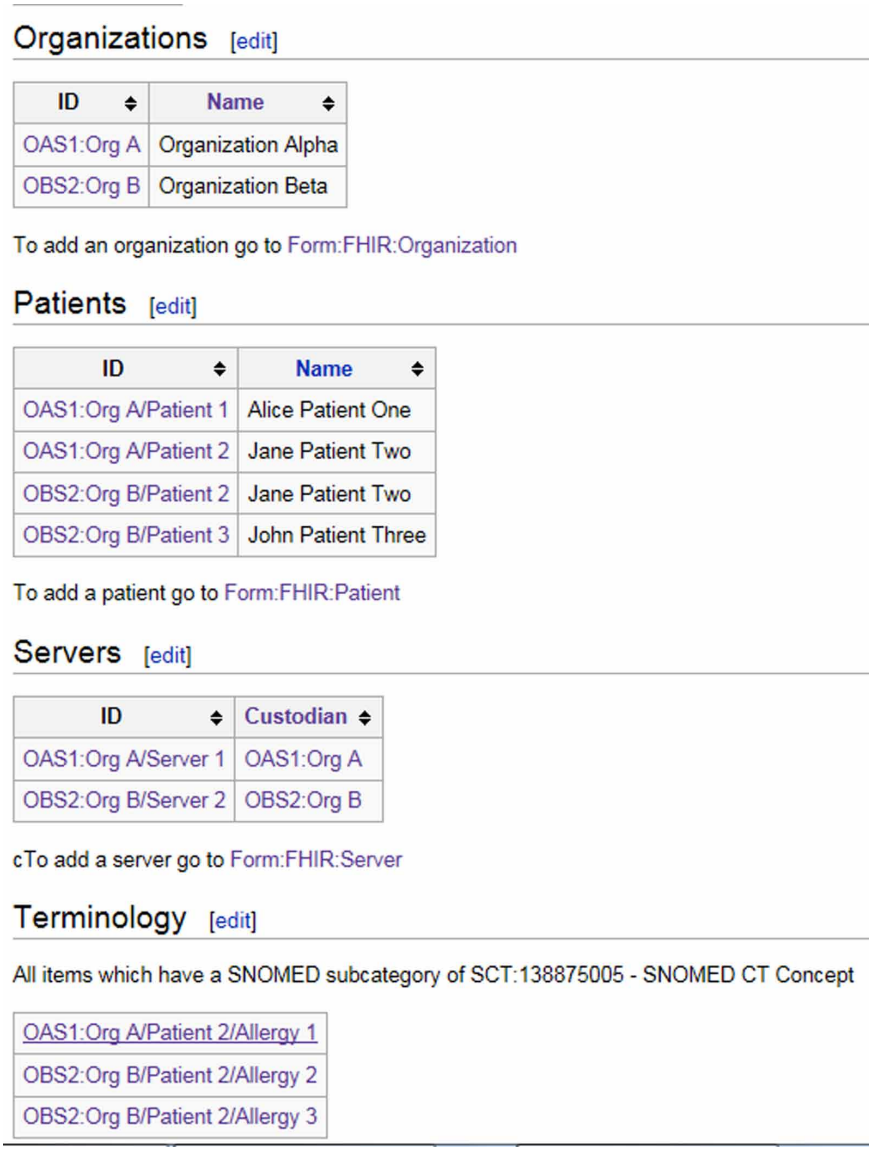
1. FHIR:AllergyIntolerance patient
2. FHIR:Organization name
3. FHIR:Patient link
4. FHIR:Patient managingOrganization
5. FHIR:Patient name

The template for FHIR:Patient defines the field "name". It also defines that the property FHIR:Patient_name has the value from the field "name". Field names are local to a template. Although this is an oversimplification of name, it shows the relationship between the template view and the underlying ontology view of named individuals (pages) and properties in the triplestore.

The line "! Patient name" defines the label to be used in display (see 10) where the patient name is displayed. Templates control the viewing behavior of a page. Forms define the data capture behavior.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. User interface view of the ontology instances



Organizations [edit]

ID	Name
OAS1:Org A	Organization Alpha
OBS2:Org B	Organization Beta

To add an organization go to Form:FHIR:Organization

Patients [edit]

ID	Name
OAS1:Org A/Patient 1	Alice Patient One
OAS1:Org A/Patient 2	Jane Patient Two
OBS2:Org B/Patient 2	Jane Patient Two
OBS2:Org B/Patient 3	John Patient Three

To add a patient go to Form:FHIR:Patient

Servers [edit]

ID	Custodian
OAS1:Org A/Server 1	OAS1:Org A
OBS2:Org B/Server 2	OBS2:Org B

cTo add a server go to Form:FHIR:Server

Terminology [edit]

All items which have a SNOMED subcategory of SCT:138875005 - SNOMED CT Concept

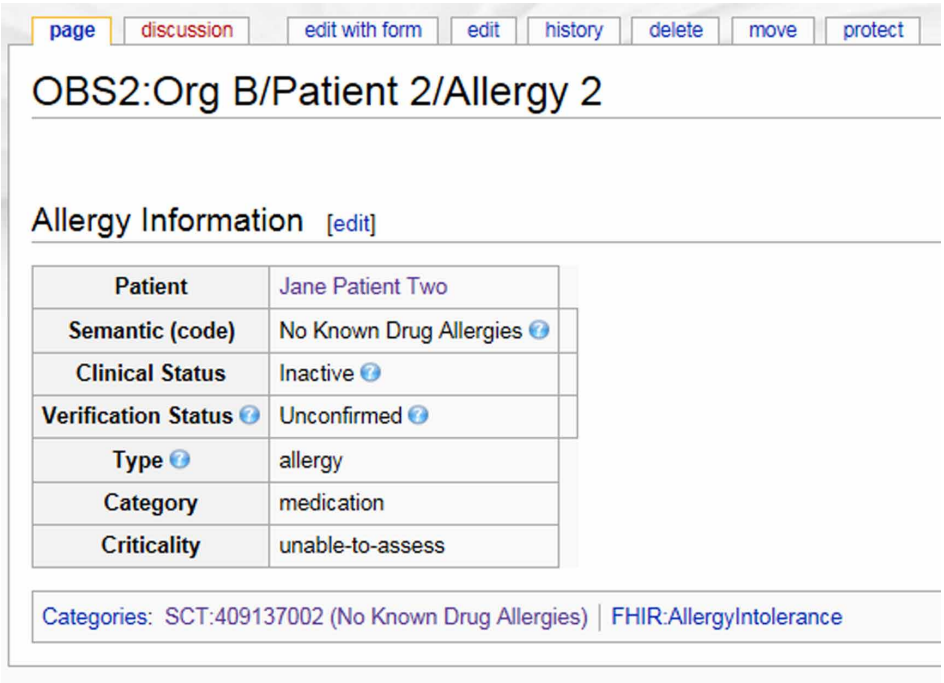
OAS1:Org A/Patient 2/Allergy 1
OBS2:Org B/Patient 2/Allergy 2
OBS2:Org B/Patient 2/Allergy 3

Code Fragment 11. (FHIR Fields in SMW template)

```
{{FHIR:Patient
|managingOrganization=
|name=
|link=
}}
..
! Patient name
| [[FHIR:Patient_name::{{{name|}}}]]
|-
```


EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. User interface view of an allergy



Allergies

Allergies, linked to the term, are displayed in the user interface in Figure 3.

To do this, the two ontologies—the patient record with the code "SCT:409137002 (No Known Drug Allergies)" and terminology categories with semantic concepts were entered into SMW. Therefore, there is closure on the concept class in SNOMED and the instances of the class in the FHIR patient record. There is no programming required to make the match.

The allergy has two categories:

1. **Semantic Type:** SCT:409137002 (no known drug allergies)
2. **Structural Type:** FHIR:AllergyIntolerance

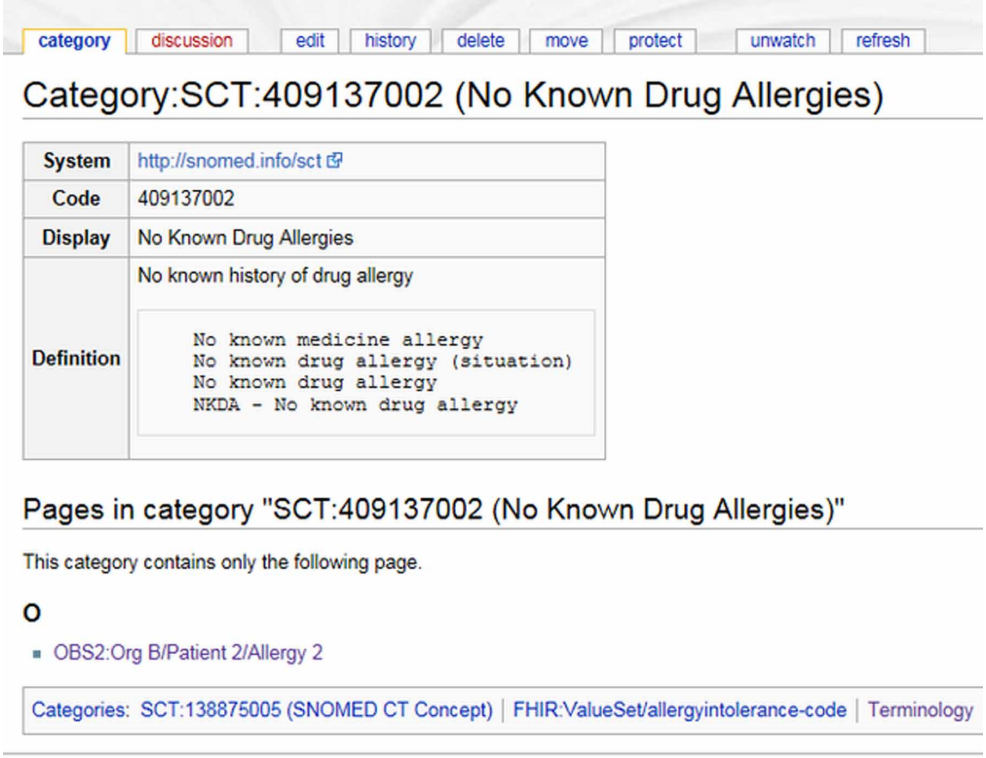
The display property is derived from the rdfs:label of SCT:409137002 (see Code Fragment 3) and can be accessed from the "Display title of" property of the concept category which is already linked. Note that SMW allows properties in Category pages. This additional information to the term allows transformation from the HL7 Codeable Concept which is expected with an implementation of the EHR with APIs supporting HL7 FHIR exchanges with other systems.

The category display provided by the platform shows its identity, label, and the super-category SCT:138875005 (SNOMED CT Concept). This is the root of the SNOMED CT ontology (see Figure 4).

For ease of implementation, the name of the concept includes its semantic in parentheses. This makes it very useful in the structured data capture forms shown later.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. User interface view of a SNOMED CT concept



The terminology becomes an integral part of the health information navigation. It allows the simple reference to the "Display title of" value which is equivalent to the rdfs:label.

Definition is used as information mouse over for an element in Figure 5.

Here is the SMW Wikitext defining the mouse over information:

The definition is retrieved from the terminology definition in the specific category – in this case "FHIR:allergy-clinical-status/inactive" which has the definition property entered (Figure 6).

Allergy Intolerance SMW Instance

The instance of the allergy intolerance in SMW conforms to the template Template:FHIR:AllergyIntolerance where the fields reference the appropriate other instance in the case of patient and category for all the terminology links.

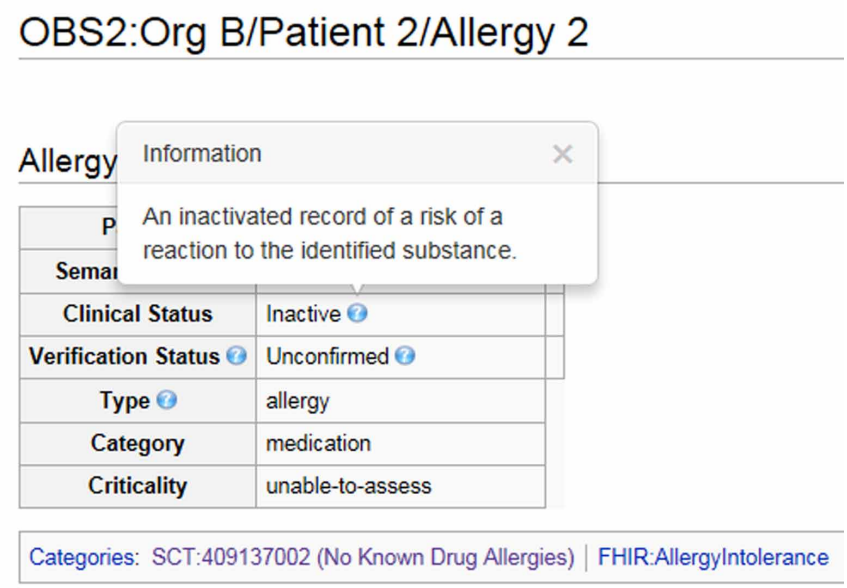
The semantic field is treated differently in that it applies to the whole Allergy Intolerance so that the instance is typed No Known Drug Allergy as seen at the bottom of Figure 3.

ValueSet Implementation

A ValueSet specifies a set of codes drawn from one or more code systems to specify which codes can be used in a context.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

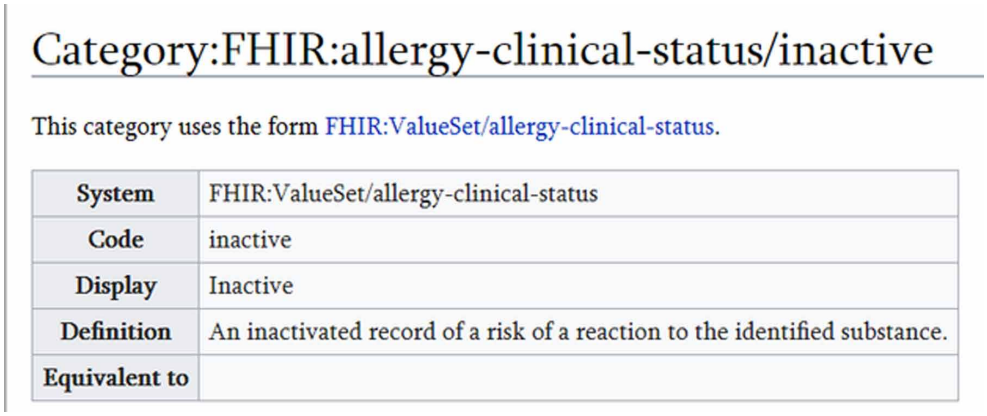
Figure 5. User interface with mouse over definition



Code Fragment 12. (SMW Template fragment for Mouse Over)

```
{{#info:{{#show:Category:{{{clinicalStatus}}}|?Terminology definition=}}}}
```

Figure 6. Clinical Status Category-inactive



Category:FHIR:allergy-clinical-status/inactive

This category uses the form FHIR:ValueSet/allergy-clinical-status.

System	FHIR:ValueSet/allergy-clinical-status
Code	inactive
Display	Inactive
Definition	An inactivated record of a risk of a reaction to the identified substance.
Equivalent to	

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Code Fragment 13. (FHIR Fields in SMW template)

```
{{FHIR:AllergyIntolerance
|patient=OBS2:Org B/Patient 2
|semantic=SCT:409137002 (No Known Drug Allergies)
|clinicalStatus=FHIR:allergy-clinical-status/inactive
|verificationStatus=FHIR:allergy-verification-status/unconfirmed
|type=FHIR:v3/NullFlavor/NA (not applicable)
|category=FHIR:v3/NullFlavor/NA (not applicable)
|criticality=FHIR:v3/NullFlavor/NA (not applicable)
}}
```

In this simplest approach, the member classes of the terminology are subClassOf a class representing the ValueSet. This example shows it as FHIR:ValueSet/allergyintolerance-code to align it with the HL7 FHIR ValueSet (see Figure 7).

The example's SNOMED concept and any SNOMED subcategories would be allowable for the AllergyIntolerance code. In theory, the assignment of FHIR:ValueSet/allergyintolerance-code would mean that the type has not been declared. Therefore, it is abstract.

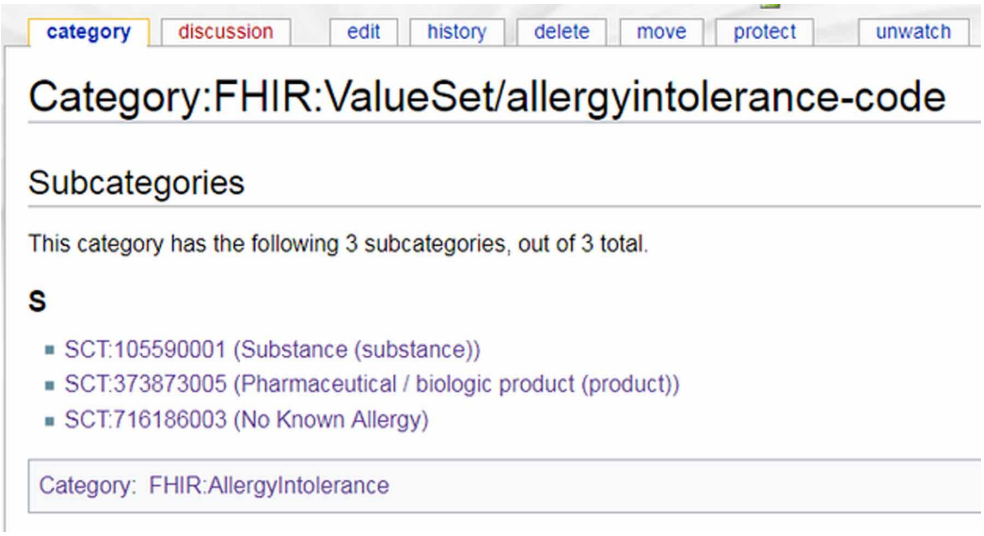
FHIR:ValueSet/allergyintolerance-code is bound to the FHIR:AllergyIntolerance category as a subcategory in SMW. Thus, the declaration of an individual as rdf:Type FHIR:AllergyIntolerance means that the semantic types are constrained to the ValueSet subcategory hierarchy. This includes the SNOMED concepts.

In UML, this ValueSet is represented with a stereotype «valueset» as shown in Figure 8.

FHIR:ValueSet/allergyintolerance-code is shown as abstract (UML italics), which means that it cannot be applied to an individual. This is interpreted later in the SMW form as hideroot in the selection.

The platform user interface for assigning the concept from the ValueSet is supported by a form construct. This provides a "tree" input type to display a category tree box (excluding abstract types). Figure 9 is an example of the form user interface during editing of Allergy 2.

Figure 7. User interface of a ValueSet



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. UML representation of the ValueSet

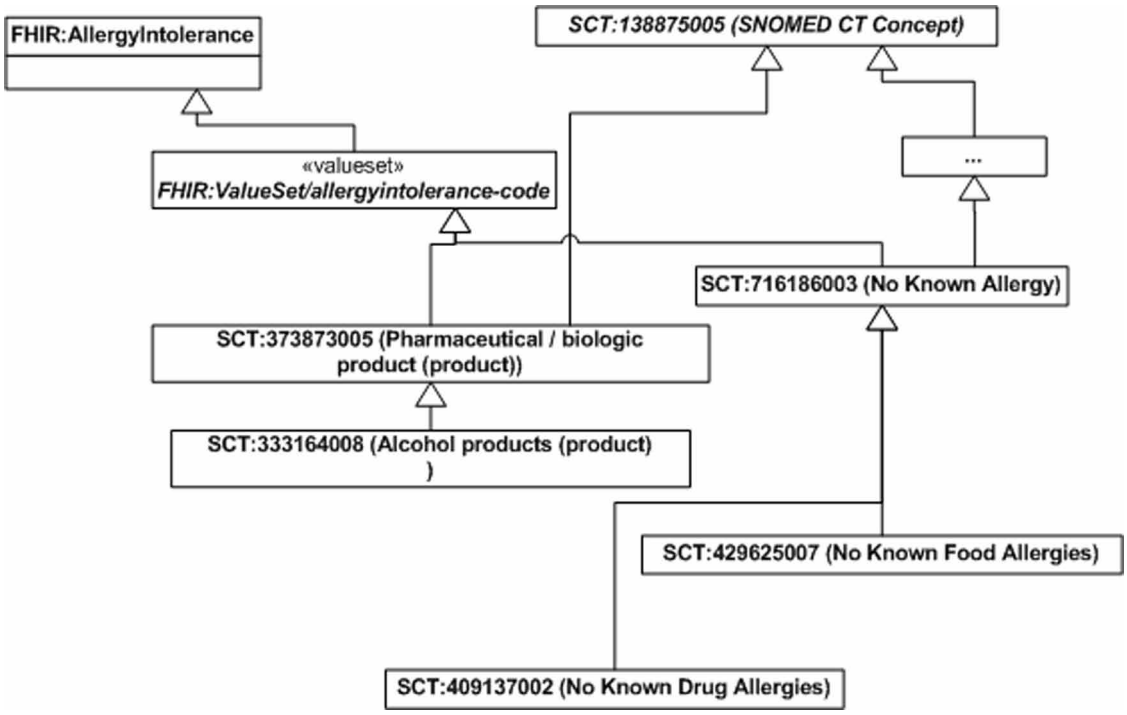
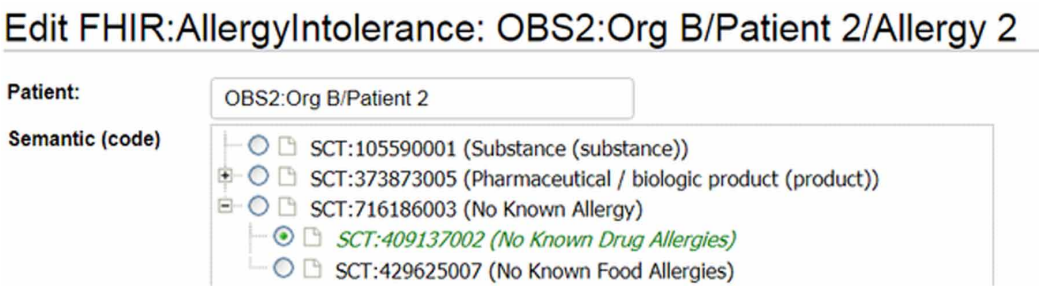


Figure 9. SMW ValueSet tree for selection in Data Capture



The Wikitext to support the semantics field entry in SMW looks like:

The top category is declared as the ValueSet. The abstract value set is hidden with hideroot. In a similar way, other fields have dropdown lists to reflect the allowable values and links.

Linked Data

If property targets in the graphs support URL dereferencing outside the wiki, then we can support a network of SMW stores where the location of the data is flexible. Although designed but not tested in the proof of concept, SMW supports an interwiki access mechanism to distribute pages.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Code Fragment 14. (SMW Form Fragment defining a Field with its ValueSet restriction)

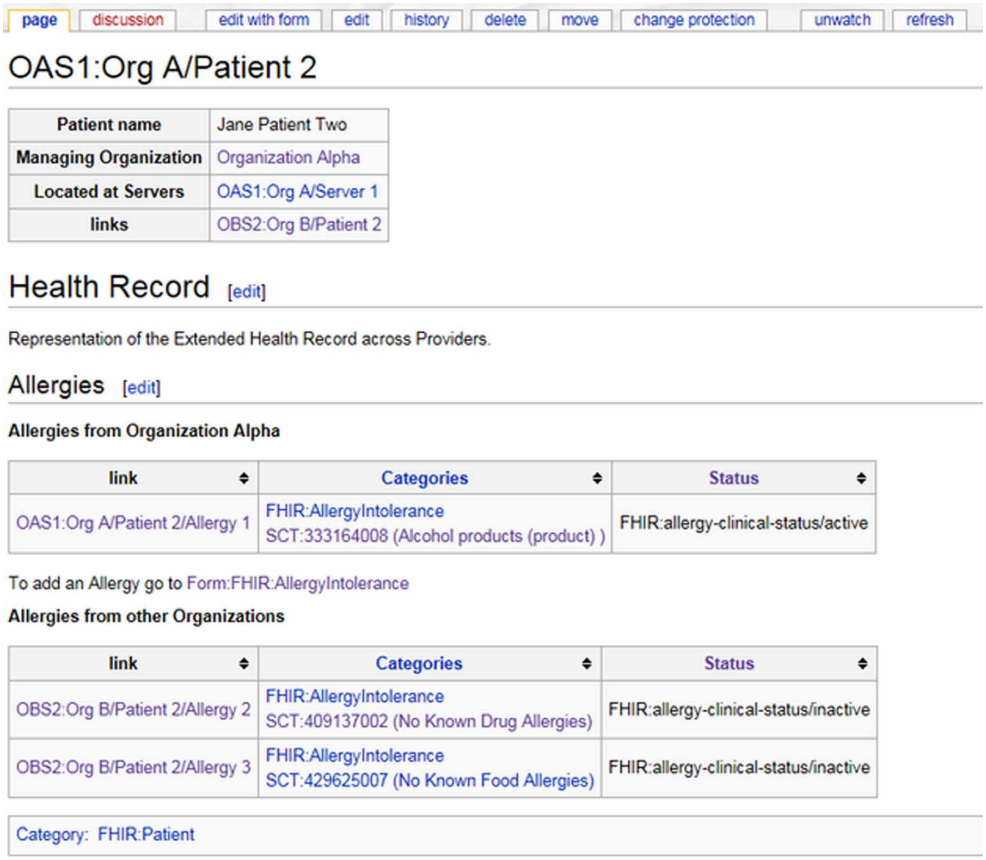
```
{{{field|semantic|input type=tree|top category=FHIR:ValueSet/allergyintolerance-code|rows=20|hideroot|}}}
```

The proof of concept explored the distributed nature of the health exchange (although it is deployed on only one wiki). Prefixes representing servers (e.g. OAS1, OAS2) can be expanded by the platform to the full dereferencable URL needed for access to that page on that external wiki (see Figure 10).

The example had "Jane Patient Two" linked across alpha and beta organizations. This view is of the patient resource at Org A on Server 1. The patient is linked to OBS2:Org B/Patient2 on Server 2. The extended health record can be displayed through in-line SPARQL-like queries supported by SMW. They walk the semantic graph from OAS1:Org A/Patient 2 to find resources assigned to the primary patient or linked patients at any known server.

The split of the display between the local context—organization alpha and other organizations—is a design decision in the prototype. The structure and formatting decisions of display are supported SMW with extensions to Wikitext specifically the #ask inline query parser function. The #ask parser function can navigate property chains including inverse properties by accessing the triplestore.

Figure 10. SMW user interface for patient with health record summary



EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Code Fragment 15. (In line query using #ask parser function)

```
'''Allergies from other Organizations'''

{{#ask:[[Category:FHIR:AllergyIntolerance]]
[[FHIR:AllergyIntolerance_patient.-FHIR:Patient_link::{{PAGENAME}}]]
|mainlabel=link
|format=table
|?Category=Categories
|?FHIR:AllergyIntolerance clinicalStatus.Display title of=Status
|headers=plain
}}
```

The query for "Allergies from Other Organizations" uses chained properties to acquire the Allergy Intolerances for the linked patient (inverse property) as well as the Display title of for clinical status.

The query result is rendered automatically by queries executing when the user navigates to a patient page. The example above is an embedded query in the Patient template page which is passed to all pages (i.e. Patients) using this template.

The compactness and power of the SMW extensions to Wikitext are reasons for the high development productivity in SMW.

Power of semantic navigation was enabled by having health record structure (e.g., HL7 FHIR) and terminology (e.g., SNOMED CT) expressed in SMW Semantic Web mechanisms of Categories, Individual Pages, and Properties and controlled by Templates and Forms.

CONCLUSION

The paradigm shift is threefold in this prototype:

1. Implementation of health record ontologies in Semantic MediaWiki with terminology references using a single identifier
2. Use of category for terminology terms defined as ontologies replacing the need for coded elements for semantics
3. Programming changes to declarative semantic annotation of Wikitext instead of traditional programming language.

Two conclusions are reached in this proof of concept:

1. The binding from a health record element to a term in a terminology is facilitated using closure mechanism in a semantic web platform.
2. Semantic MediaWiki is a feasible platform for implementing an EHR with binding to terminology which is loaded onto the platform. Semantic annotation is powerful enough to implement an EHR without recourse to traditional programming languages.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Binding Conclusion

Direct mapping to type in a Semantic Web platform allows the bindings to occur with an instance to semantic category (class) link in addition to instance to structural class. Expression of the terminology as an ontology of classes and subclasses is supported by behavior and integration of terminology into the complete picture of the executing health record.

The proof of concept was to implement this binding in an open data platform, specifically Semantic MediaWiki. The proof of concept has been demonstrated to work and screen shots are included in this paper from the working prototype.

Platform Conclusion

In implementing a proof of concept in Semantic MediaWiki (SMW), the ease and velocity of implementation were impressive. The entire initial prototype took less than 10 hours to implement. Additional user interface refinement and additional elements to the fragments extended this time.

While SMW is not thought of for implementing mission critical applications, it will eventually be adopted for this purpose. It is an example of potential sixth generation computing, which combines high-level language and Semantic Web into a platform for application development and execution.

In this paradigm, a model-driven design, system distribution, and interoperability are built into the platform rather than the application. The SMW platform used in the example is not specific to healthcare.

This approach is not in conflict with HL7 FHIR because it requires a unified platform whereas FHIR can bridge between platforms. It would be expected that a SMW implementation would provide FHIR APIs to connect with external systems which do not use SMW.

REFERENCES

Gartner Inc. (2003). *OlivaNova Benchmark, request for examination to care-technologies @care-t.com.* Retrieved from http://www.gartner.com

HL7. (2018). *valueset-allergyintolerance-code*. Retrieved from http://hl7.org/fhir/valueset-allergyintolerance-code.html

HL7 FHIR. (2017). *FHIR:Allergyntolerance*. Retrieved from http://www.hl7.org/fhir/allergyintolerance.html

HL7 FHIR. (2018). *Welcome to FHIR*. Retrieved from http://hl7.org/fhir/

HL7 FHIR/RDF. (2017). *RDF_for_Semantic_Interoperability*. Retrieved from http://wiki.hl7.org/index.php?title=RDF_for_Semantic_Interoperability

HL7 Vocabulary. (2018). *ValueSets*. Retrieved from http://hl7.org/fhir/valueset.html

International Health Terminology Standards Development Organization (IHTSDO). (2014). *SNOMED CT URI Standard v0.8.*

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Kamel Boulos, M. N. (2009). Semantic Wikis: A Comprehensive Introduction with Examples from the Health Sciences. *Journal of Emerging Technologies in Web Intelligence*, *1*(1), 94–96.

LOINC. (n.d.). Retrieved from https://loinc.org/

Marco-Ruiz, L., Pedrinaci, C., Maldonado, J. A., Panziera, L., Chen, R., & Bellika, J. G. (2017). Publication, discovery and interoperability of Clinical Decision Support Systems: A Linked Data approach. *Journal of Biomedical Informatics*, 62, 243–264. PMID:27401856

Object Management Group. (n.d.).

OWL API. (n.d.). Retrieved from http://owlcs.github.io/owlapi/

Rector, A., Qamar, R., & Marley, T. (2009). Binding Ontologies & Coding systems to Electronic Health Records and Messages. *Journal of Applied Ontology, (1)*, 51-69.

Semantic MediaWiki Project. (2017). *Main Page*. Retrieved from https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki

SNOMED International. (n.d.). Retrieved from https://www.snomed.org/

Solbrig, H. (2008, 12 9). Semantic mediawiki. Retrieved from http://nkos.slis.kent.edu/2008workshop/HaroldSolbrig.pdf

World Health Organization. (n.d.). ICD. Retrieved from http://www.who.int/classifications/icd/en/

KEY TERMS AND DEFINITIONS

4GL: A fourth generation (programming) language (4GL) is a grouping of programming languages that attempt to get closer than 3GLs to human language, form of thinking and conceptualization. 4GLs are designed to reduce the overall time, effort and cost of software development. The main domains and families of 4GLs are: database queries, report generators, data manipulation, analysis and reporting, screen painters and generators, GUI creators, mathematical optimization, web development and general-purpose languages.

Clinical Terminology: Clinical terminology is a structured vocabulary used in clinical practice to accurately describe the care and treatment of patients. It covers complex concepts such as diseases, operations, treatments and medicines.

FHIR: Fast Healthcare Interoperability Resources (FHIR, pronounced fire) is a standard describing data formats and elements (known as resources) and an Application Programming Interface (API) for exchanging Electronic Health Records.

Semantic MediaWiki: Semantic MediaWiki (SMW) is a free, open-source extension to MediaWiki – the wiki software that powers Wikipedia – that lets you store and query data within the wiki's pages. Semantic MediaWiki is also a full-fledged framework, in conjunction with many spinoff extensions, that can turn a wiki into a powerful and flexible knowledge management system. All data created within SMW can easily be published via the Semantic Web, allowing other systems to use this data seamlessly.

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Semantic Web: According to the W3C (World Wide Web Consortium), "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries". The term was coined by Tim Berners-Lee for a web of data that can be processed by machines.

ValueSet: A ValueSet is a list of codes and corresponding terms, from NLM-hosted standard clinical vocabularies (such as SNOMED CT, RxNorm, LOINC and others), that define clinical concepts to support effective and interoperable health information exchange.

Wiki: A wiki is a website on which users collaboratively modify content and structure directly from the web browser. In a typical wiki, text is written using a simplified markup language and often edited with the help of a rich-text editor.

Wikitext: Wikitext, also known as Wiki markup or Wikicode, consists of the syntax and keywords used by the MediaWiki software to format a page.

This research was previously published in the International Journal of Privacy and Health Information Management (IJPHIM), 6(2); edited by Francesco Longo and Letizia Nicoletti; pages 50-71; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Index

4GL 2048-2049, 2057, 2070

A

Ab Initio 505, 1083, 1173, 1558-1559, 1577
Abantu Illnesses 986-987, 992, 994-995, 1002, 1004-1005
Access Control 134, 169, 171, 173-174, 177, 179, 182-190, 451, 458-459, 465-466, 546, 593, 597-600, 607-610, 1227, 1255, 1273-1274, 1371, 1378-1384, 1386, 1390-1391, 1394-1396, 1406, 1485-1490, 1495, 1498-1499, 1711, 2003
Access Control Models 182-183, 185-190, 1487
Access Control Policies 187, 1378, 1381, 1384, 1390-1391, 1394-1396, 1485-1489, 1495, 1498-1499
Accountable Care Organization 1860, 1945
Action Recognition 1006-1007, 1009, 1017
Activity Theory 1-7, 9
Adjunct Mobile Support 879
Adventitious Respiratory Sounds 815-816, 821, 827, 829
ALCHEMIST 369-370, 377-379, 382, 384
Annotation 107, 109, 331, 334-335, 473-474, 529, 554, 556-557, 559, 563-565, 589, 817-819, 822-823, 825, 828, 1053, 1168-1174, 1177-1180, 1183-1184, 1187, 1192, 1508, 1514, 1559, 1816-1817, 1966, 1978, 2057
ANOVA 582, 585-586, 588-589, 885, 1655
Artificial Intelligence 93-94, 390-395, 397-399, 416, 472, 513, 666, 686, 869, 933, 1461, 1521, 1971, 1993, 1995, 2001, 2049
Artificial Neural Network (ANN) 50, 54, 58, 292, 391-395, 397, 487-488, 490, 503, 527, 676, 766, 834, 836, 844, 1827, 1895, 1919, 1994
Assisted Living 1606, 1611, 1963, 1977
ATP 381, 1172, 1879, 1894
Attack Scenario 1797-1799, 1802, 1811
Attributional Theory 986, 988
Automatic Detection 815-816, 818, 827

B

Backpropagation 487, 490, 492, 494
Bacterial Foraging 565, 572, 780
Bamidis 1602, 1609, 1966, 1969, 1972, 1977
BAN Logic and Scyther Tool 439
Base Learners 654-655, 658-659, 707, 709, 711, 713, 717-718, 722, 725
Batch Correction 580, 582, 584-585, 589
Bayesian Network 55, 227, 244, 508, 562, 1804-1805, 2005
Bdellovibrio 908, 910, 928-929
Biclustering 1306, 1310-1313, 1318, 1321-1322, 1325, 1327
Bidirectional Search 471, 476, 478, 481, 1814, 1819, 1822
Big Data 31, 51-52, 55-57, 74, 76, 83-85, 90, 94, 155, 298, 411-412, 421, 608, 610, 869, 1154, 1162, 1197, 1200, 1209, 1365-1366, 1401-1402, 1404, 1407-1410, 1412, 1523-1524, 1541, 1760, 1826-1828, 1831, 1836, 1839-1852, 1989-1992, 1994-1997, 2000-2001, 2035, 2042
Big Data Analytics 83, 85, 1154, 1401, 1412, 1760, 1839-1841, 1843, 1850
Binding Mode 1560, 1567, 1577
Biodegradation 1557, 1568
Bioinformatics 77, 206, 298-299, 319-320, 472, 502-503, 514, 520, 523, 554, 562, 565, 580-581, 629, 631, 633, 639, 659, 708, 1024, 1052-1053, 1171, 1306-1307, 1310, 1366, 1513, 1557, 1833, 1875, 1877-1878, 1881-1882, 1899, 1909, 1911, 1994, 1996
Bio-Inspired Algorithms 1898-1899, 1905-1906, 1909-1911, 1919
Bio-Inspired Computing 523, 1306, 1327
Biomarkers 410-411, 420, 551, 563, 653, 1875-1878, 1894, 1965, 1967, 1972, 1975, 1977
Bio-Med 581-582, 589

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Biomedical Applications 249-250, 265, 852, 860, 1076-1077, 1079, 1100, 1882
Biomedical Health System 987, 991-993, 996, 998, 1001-1002
Biomedical Informatics 59, 1370
Biomedical Science 1283-1284, 1286, 1291, 2002
Biometrics 31, 1256, 1259-1261, 1264, 1267, 1272-1273, 1275, 1282
Bio-Nano-Informatics 1875, 1894
Biosensor 672
Blood Glucose Levels 13-19, 24-27, 30-34, 36, 38-40
Body Area Networks 270, 459, 971, 973, 980, 1404, 1782, 1797
Body Sensor Network 457, 663
Bolus 13, 22, 25-26, 30, 33-34, 39
Boolean Networks 226-228, 235, 244
Bottom-Up 76, 466, 562-563, 898-899, 902, 1705, 1740
Box Filtering 815, 818, 821
Box Plots 581-582, 586, 588-589
Brain Computer Interface 842
Breast Cancer 58, 113, 210, 376, 779, 782, 794-795, 800, 808, 1058-1059, 1228, 1328-1330, 1337, 1367, 1531, 1647, 1877-1878
Business Intelligence (BI) 872, 1759-1760, 1777
Business Process Management (BPM) 136, 1739

C

Cancer Therapy 203
Cardiovascular Disease 16, 31, 34-36, 208, 664, 670, 705, 1225, 1611
Care Providers 79-80, 170, 172, 465, 531-532, 1198, 1237, 1298, 1416-1418, 1421-1428, 1431, 1433-1437, 1442-1445, 1451, 1764, 1766, 1859, 1927, 1947, 1964
CARMA 1877, 1894
Cellular Potts Model 388
Central Dogma 504, 1169, 1904, 1919
Chirography 1263, 1282
Classification Algorithms 680, 765
Cleavage Cycle 373, 381, 383-384, 388
Clinical Decision Support 390-391, 393-394, 396-399, 496, 828-829, 868, 1264, 1461, 1583, 1711, 1846, 1946, 1948, 1955, 1990, 2020, 2022, 2050
Clinical Terminology 532, 2070
Cloud Computing 74, 76-77, 83, 94, 137, 437, 593-594, 596, 608-610, 1024, 1054, 1065, 1259, 1266, 1364-1366, 1368-1374, 1411, 1523, 1787, 1808, 1864, 1922-1923, 1932-1933, 2000, 2020
Cloud Services 78, 532, 598, 609, 1255, 1259, 1366, 1371, 1373, 1523, 1782, 1995-1996
Clustering 53, 58, 473, 503, 507, 512, 517, 555-556, 585, 789, 1060, 1064, 1169-1170, 1306, 1310-1311, 1313, 1316-1317, 1321-1322, 1325, 1327, 1816, 1827, 1877, 1995
CMOS 852-853, 856, 860
Coding Regions 506, 1168, 1172
Cognitive Decline 1964, 1969-1970, 1974-1975, 1978
Collaborative Healthcare Information Systems 1-2, 4, 7
Collective Behaviour 502, 513-514, 523-524
Collective Intelligence (CI) 1759, 1761
Community Cloud (CC) 437, 452
Component Analysis 128, 202, 517, 582, 584, 711, 713, 725, 772, 780, 1013, 1053-1058, 1064, 1066, 1071, 1312, 1587, 1655, 1658, 1827
Compound Comparison 629, 631, 633
Compression Ratio 463, 672, 931-934, 938-939
Computational Models 250, 369, 373, 561, 925
Concept Attribute Labeling 325, 334, 336-337
Conceptualization 137, 173, 1211, 1241, 2070
Conditional Random Field (CRF) 1509, 1520
Consent Ontology 1241-1242, 1246, 1253
Context-Aware 188, 269-271, 325, 327, 459, 1149, 1964, 1970
Continuous User Authentication 1282
Correlation-Based 716, 725
Crackles 815-825, 827-829
CUDA 629-636, 638, 1059, 1066, 1068, 1833
Cultural Modernity 1005
Cytopathology 1921-1931, 1933-1934, 1942
Cytoscape 556-557, 900

D

Data and Information Quality (DIQ) 1758-1759
Data Compression 463-464, 466, 669, 672-673, 930-933, 938-941, 944, 950, 952, 955-956, 959, 964, 967, 981
Data Extraction 425, 427, 1840
Data Mining 31, 50-69, 398, 410, 412, 421, 432, 646, 872, 974, 979, 1025, 1306, 1310, 1313, 1741-1742, 1762, 1777, 1799, 1831, 1877, 1971, 1978, 1989, 1993-1997, 2000
Data Protection 91, 819, 1201, 1460, 1582, 1840
Data Visualization 865, 872, 1066, 1148, 1150, 1154-1156, 1162, 1166, 1741, 1743, 1745, 1750
Data Work 1148, 1150-1153, 1156, 1166
De Brujin Graph 1173-1174, 1177, 1186-1187, 1191
De Novo 1172, 1181, 1186, 1558, 1577
Decision Making 50, 126, 129, 150, 341, 390-391, 393-394, 398-399, 473, 678, 858-860, 862-863, 981, 1151, 1154, 1162, 1451, 1579-1580, 1583,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1590-1591, 1614, 1742, 1758, 1760-1761, 1763-1764, 1777, 1816, 1840, 1955, 1965, 1992-1993, 2000, 2020, 2023
Decision Support 67, 343, 390-391, 393-394, 396-399, 481, 496, 746, 828-829, 868, 933, 1196, 1241, 1264, 1418, 1422, 1461, 1583, 1705, 1710-1711, 1714, 1728, 1760, 1803-1804, 1846, 1946, 1948, 1955, 1965, 1970-1971, 1979, 1990, 2020, 2022, 2050
Decision Tree 54, 58, 417, 713, 716-718, 722, 727, 780, 1508, 1994
Deep Learning 337, 728-729, 1290, 1826-1828, 1830-1834, 1836
Developmental Biology 369-370, 372, 379, 382
Diabetes 13-41, 58, 77, 270, 331, 456, 700, 973, 995, 1023-1024, 1031-1032, 1610-1611, 1812, 1922, 1954, 1995, 2004
Diagnostic Imaging Examinations 1522-1524, 1528, 1530-1532
Diffusion of innovation 1416, 1423, 1428, 1443
Digital Crimes 1781-1782, 1788, 1794-1795, 1807-1809, 1811
Digital Forensic Investigation 1812
Discrete Wavelet Transform (DWT) 687, 747, 963, 967
Disease Ontology 473-476, 481, 532, 540, 1814, 1817, 1822
Distributed Computing 1054, 1403, 1832, 1842, 1845, 1849, 1996
DNA Computing 1879, 1882-1883
DNA Fragmentation 1192
Drug Discovery 195-197, 199, 201, 207, 209-211, 563, 1910
Dynamic Programming 511, 565, 691, 1183

E

ECG Compression 930-932, 937-939, 967-968
E-Health 69, 86, 92, 135-136, 149, 156, 357, 359-360, 363, 366, 459, 663, 699-700, 705, 1201, 1210, 1383, 1404-1405, 1407, 1410, 1419-1422, 1448, 1454-1455, 1458, 1461, 1581, 1599, 1633-1634, 1640, 1642, 1782-1785, 1787, 1789-1791, 1793-1795, 1804, 1807, 1812, 1843, 1942, 1952
E-Health Systems 1782-1784, 1789, 1791, 1794-1795, 1807, 1812
EHR 135-136, 170-177, 325-330, 332, 334, 336, 391, 393, 395, 397, 411, 429, 464-466, 496, 529, 531, 533-534, 542, 593, 614-619, 623-624, 665, 1039, 1258, 1264-1267, 1269, 1275, 1358, 1387, 1392, 1417, 1451-1453, 1455-1457, 1459-1461, 1467-1470, 1473, 1475-1482, 1485-1489, 1492-1494, 1496, 1506, 1581, 1763, 1787, 1791, 1795, 1846, 1849-1850, 1858, 1860, 1866, 1868-1871, 1943, 1945, 1947-1951, 1954, 1956, 1991, 2049, 2051-2052, 2059, 2062
E-Learning 609, 1942
Electrocardiogram (ECG) 666, 744-745, 747, 763, 930-931, 1404
Electroencephalography (EEG) 1654-1655
Electromyogram 762-763, 834, 838, 840, 931, 975, 1800
Electromyography 762, 833, 835-837, 1404, 2007
Electroneurography 835, 842
Electronic Health Records (EHR) 75, 120, 129, 135-136, 155, 172, 325-327, 329, 391, 411, 429, 465, 528-529, 531, 593, 609, 614-620, 625, 665, 699, 1039, 1041, 1162, 1196, 1227, 1258, 1264, 1364, 1366, 1373, 1387-1388, 1416-1423, 1425, 1428, 1437-1438, 1442-1444, 1449, 1451-1452, 1467-1470, 1485-1486, 1581, 1722, 1763, 1787, 1791, 1795, 1846, 1849-1850, 1858-1860, 1943, 1945, 1947-1948, 1956, 1990-1991, 2049, 2059, 2070
Electronic Medical Record (EMR) System 80, 117-123, 125, 127-130, 347, 528-529, 531, 614, 626, 978, 1039, 1199, 1201, 1222-1223, 1235, 1253, 1255-1256, 1258, 1297-1298, 1450, 1583, 1674, 1722-1723, 1726, 1788, 1791, 1923, 1943, 1948, 1951
Electronic Records 7, 125-126, 529, 663, 1298, 1419, 1428, 1578-1580, 1582-1583, 1591, 1846, 2018
Electronic Records Management System 1578-1580, 1591
Embedded Link System 667-668, 705
Emergence 210, 374, 514, 899, 1007, 1017, 1194, 1254, 1258, 1407, 1499, 1539, 1683, 1699, 1747, 1827, 1859, 1943-1944, 1954, 1956, 1969
Emerging Technologies 1271, 1931
EMR Adoption Model 1455
EMR System 118-119, 129, 1222, 1224-1226, 1231-1233, 1236, 1242-1243, 1246-1247, 1297-1298, 1301-1303, 1459, 1674
Ensemble Learning 645-646, 651, 658, 662, 707-709, 718, 725
Ensemble Method 655, 662, 709, 711
Entropy 206, 208, 710-711, 728, 764, 932, 1056-1057, 1328-1331, 1334-1335, 1337, 1505, 1513
Enzymes 18, 553, 559, 1169, 1556-1558, 1561, 1565, 1568, 1908, 1919
Euler Path 1173-1174, 1177-1180, 1186-1187, 1192
Evidence Based Medicine 57, 165
Evidence-Based Design 1155-1156, 1163, 1166
Evolutionary Algorithms 1560, 1905
exergames 1599-1607, 1609-1610, 1612, 1617-1620,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1966, 1968-1970
Exons 506, 511, 1172, 1183-1185, 1188, 1192
Extreme Learning Machine 779, 781, 789, 791, 795
Extreme Programming 1341-1342

F

Fall Detection 726-730, 734-736, 739
Fanmod 558
Fast Fourier Transform (FFT) 1655, 1658
Fast Healthcare 425, 428, 430, 531, 1486, 2049, 2051, 2070
Feature Redundancy (FR) 725
Feature Relevancy 709, 715, 718
Feature Selection 413-415, 418, 421, 512, 518, 645-646, 650-653, 657-659, 662, 707-710, 714-718, 722, 780-781, 783, 788, 795, 808, 1007, 1877-1878
FHIR 424-425, 428, 430-432, 434, 529-531, 534-536, 543-545, 549, 1486, 2048-2049, 2051, 2053, 2056, 2059-2060, 2062-2063, 2065, 2068-2070
Filtering 104, 201, 291, 350, 462-463, 558, 668, 689, 764, 815, 818, 820-821, 1026, 1069, 1071, 1166, 1411, 1472, 1487, 1661, 1901
Formal Techniques 1406, 1793
Fractal Dimension 815, 818, 820-821
Functional Enrichment 554, 556, 563, 1062
Fuzzy Logic 50, 55, 285, 394, 503, 510, 676, 766, 1312
Fuzzy System 394, 523, 1312
Fuzzy-Rough Feature Selection (FRFS) 662

G

Gait 762-764, 767-775, 837-839, 845, 1966, 2003, 2008
gamified applications 1600, 1607, 1609-1610, 1618-1620
Gene Annotation 1168, 1171-1174, 1177-1179, 1187, 1192
Gene Assembly 1192
Gene Expression 226, 244, 371-372, 381, 388, 502, 507-509, 512-513, 517-518, 553, 556, 580-581, 583, 589, 645-647, 652, 658, 715, 779-781, 1052-1053, 1055-1061, 1063-1064, 1066, 1068-1069, 1306-1307, 1829, 1876-1878, 1880, 1894
Gene Ontology 554, 1877
Gene Regulatory Network 374, 380, 507, 1052-1061, 1064, 1066, 1068-1069, 1071, 1908
Genetic Algorithm (GA) 50, 54-55, 202, 391-395, 397, 503, 568, 571, 647, 650, 662, 743-744, 749-753, 758, 780, 1312, 1560, 1895, 1905, 1919
Genetic Counseling 1226, 1236-1238, 1240, 1253
Genetic Engineering 1911, 1919
Genetic Services 1222, 1224, 1233-1237, 1241, 1253
Genome 77, 298, 319-320, 506, 508, 511, 551, 554, 558-559, 562, 586, 1053, 1063, 1169-1172, 1174, 1178, 1180, 1182-1184, 1186-1188, 1192, 1367, 1374, 1817, 1847, 1909
GEO 1877, 1894
GFSK Demodulator 852-854, 860-861
GIS 1522-1524, 1530, 1532
Glycaemic Control 14-15, 17, 19, 28, 30, 34, 36, 40
gm-C Filter 856
GO 13, 25, 27, 39, 327, 391, 459, 514, 531, 554, 556, 595, 686, 743-744, 915-916, 918, 997, 1152, 1156, 1259, 1269, 1385, 1389, 1403, 1423, 1482, 1605-1606, 1707, 1715-1716, 1766, 1798, 1841, 1850, 1869, 1877, 1894, 1896, 1904-1905, 1945, 1965, 1970

H

Hadoop 83-84, 595, 598, 602, 606-608, 610, 1411, 1831-1832, 1840, 1844, 1996, 2001
Hashing 298, 301, 304-305, 320
Hba1c 14, 16-17, 30-31, 37, 40, 415, 1954
Health Care 1-2, 4, 6-7, 9, 31, 50-52, 55-69, 75, 79-85, 90, 117-123, 125-126, 128-130, 133-137, 139, 142-144, 149-151, 155-158, 161-165, 169-173, 177, 179, 182-190, 268-272, 275, 284-285, 289, 326-328, 336, 340-345, 349-351, 357-359, 361-362, 365, 390-394, 397-398, 410-412, 417, 424-434, 437-439, 442, 448, 451-452, 456-457, 459-460, 462, 465, 467, 472, 487-488, 496-497, 529-533, 543, 545-546, 593-595, 608-610, 615, 663-667, 671, 673, 675, 683, 686, 688, 694-696, 699-700, 705, 746, 865-869, 873, 875-876, 931, 971-972, 974-975, 978, 981, 998, 1002, 1039, 1041-1042, 1148, 1150-1153, 1155-1157, 1162, 1193-1203, 1205-1207, 1209-1211, 1221-1228, 1234, 1240-1241, 1254-1259, 1263-1264, 1266-1269, 1275, 1282, 1297-1299, 1301-1303, 1341-1342, 1345, 1361, 1364, 1366-1367, 1369, 1372-1374, 1378-1379, 1381, 1383, 1387, 1394-1397, 1401-1412, 1416-1423, 1427, 1442-1443, 1449, 1451-1455, 1457-1461, 1467-1470, 1472-1473, 1475-1482, 1485-1487, 1489, 1492-1494, 1498-1499, 1507, 1514, 1522-1524, 1528, 1530-1532, 1581, 1600, 1618, 1632-1640, 1642, 1644, 1647-1649, 1683-1688, 1690-1692, 1697, 1700-1702, 1704-1709, 1714-1715, 1720-1723, 1729-1730, 1739-1741, 1754, 1758-1766, 1768-1771, 1774-1777, 1781-1783, 1785-1790, 1792, 1794-1797, 1804, 1807-1809, 1812, 1826, 1828,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1830, 1833-1834, 1836, 1839-1845, 1849-1850, 1852, 1858-1862, 1864-1870, 1922, 1927, 1929, 1942-1943, 1945-1952, 1954-1956, 1964, 1978, 1989-1997, 2002, 2008, 2012, 2018-2024, 2033-2035, 2038-2039, 2041, 2049, 2051, 2069-2070
Health Care Reform 1943, 1956
Health Care System 118, 688, 1039, 1421, 1451, 1830, 1833, 1836, 1955-1956
Health Informatics 31, 155, 171-172, 472, 593-594, 609, 1006, 1009-1010, 1018, 1020-1021, 1162, 1201, 1301, 1690, 1815, 1833
Health Information Systems 411, 1041, 1198, 1457
Health Information Technology (HIT) Systems 1255, 1257, 1282
Health Information Technology For Economic And Clinical Health Act 411, 1473, 1739
Health Insurance Portability And Accountability Act (HIPAA) 64, 183, 438, 595, 1226, 1253, 1255, 1258, 1477, 1503, 1520, 1950
Healthcare Big Data 51, 1989, 1995-1997
Healthcare ICT Projects 2022, 2024
Healthcare Industry 79-80, 150-151, 155-158, 161-162, 165, 350, 437-438, 1222-1223, 1257-1258, 1342, 1364, 1468, 1476-1477, 1482, 1720, 1774-1777, 1843, 1858, 1866, 1868-1870, 1989-1990, 1992, 1994, 1996-1997
Healthcare Informatics 50, 58-59, 61, 65, 183, 187
Healthcare Information Systems (HIS) 133, 182, 1862
Healthcare Knowledge 59, 1632-1633, 1636, 1640, 1642, 1644, 1648
Healthcare Provider 81-82, 183, 187, 460, 595, 1151, 1449, 1451, 1467, 1759, 1787
Healthcare Services 85, 135, 149, 155-156, 184, 269-270, 275, 284, 343, 437-438, 457, 664-665, 699, 867, 1151, 1194, 1199, 1209, 1406, 1449, 1523-1524, 1528, 1531-1532, 1632, 1634-1636, 1638, 1705, 1707-1708, 1759, 1783, 1787-1788, 1794, 1809, 1952, 1992, 1997, 2020, 2023-2024, 2034
Healthcare Workflows 1378
HealthConnect 1417, 1420-1422, 1425-1427, 1444, 1449
Healthy 3.0 1193, 1209-1210, 1220
Healthy Aging 1600, 1974-1975
Heart Sound 666, 687, 693, 705
High Performance 520, 629-630, 637, 829, 1069, 1878, 1996-1997, 2000
Higher Education 482, 925, 1283, 1579
High-Throughput 195, 551-552, 557, 562, 580-582, 1876
HIS Integration 135, 144, 149
HIS Interoperability 149
HIS Measurement 1220
Holter Machine 744-745
Hospital Information Systems (HIS) 427, 866, 1450, 1454, 1583, 1674, 1720-1722, 1867, 2020
Human Research Ethics 1537-1538, 1543
Human-Data Interaction 1148-1149, 1166
Hybrid Ensemble 662, 707, 709, 716, 718, 725
Hybrid Ensemble Learning 662, 707, 709
Hyperglycaemia 13-16, 25, 32
Hypoglycaemia 13-14, 18, 26, 32, 34, 37-38

I

IBM 84, 373, 885, 1402, 1841, 1989, 1996, 2001
ICT Acceptance 1193, 1198, 1220
ICT Induce Benefits 1221
ICT Performance 1221
ICT ROI 1221
ICT Support 1221, 1711
IDC 1721, 1989
Identity Management 1405
IHE 135, 359, 363-366, 425-426, 429, 431, 433-434
IHE Profiles 359, 363, 434
Implantable Medical Devices 971, 1782, 1786, 1790, 1799, 1812
Improved Patient Workflow 1739
In Silico 200, 203, 208, 210, 552, 1558
In Vivo 195, 210, 340, 345, 1876, 1879-1881, 1883, 1894, 1909, 2037-2039, 2041
Independent Component Analysis (ICA) 1056, 1066, 1071, 1655, 1658
Indigenised 991, 995, 1005
Indigenous 990-991, 994, 996-998, 1002, 1005, 1538
Indigenous Medicine 990, 994, 996-997
Induced Fields 249-251
inertial sensor 762
Infant Cries 289, 292
Inference Engine 540, 670-672, 695, 706, 1801
Information Extraction 326, 424, 430, 1507
Information Infrastructure (II) 1706
Information Retrieval 103, 471-472, 1310, 1815
Information Security Standards 1683-1684, 1702
Information Systems (IS) 1739
Information Technology 61-62, 85, 120, 127, 136, 152, 157, 183, 187, 359, 390, 411, 530, 545, 868, 1039, 1053, 1159, 1195, 1198-1199, 1201-1202, 1223, 1255, 1257, 1275, 1282, 1298, 1347, 1351, 1365, 1373, 1379, 1418, 1450, 1467-1470, 1473, 1480, 1503, 1539, 1581, 1590, 1635, 1720, 1739, 1858-1860, 1894, 1943-1945, 1947, 1951, 1954-1955, 2022, 2034-2035

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Inherence Factor 1282
Insulin 13-15, 17-20, 22, 25-34, 36-40, 1791, 1812, 2005, 2007
Integrated Development Environment (IDE) 904, 929
Integrated Knowledge 164, 350
Interactive Maps 1522, 1524, 1530, 1532
Interoperability 85, 117-118, 120, 128-130, 134-135, 137, 149, 163-164, 169-173, 179, 186, 271, 357-360, 362-363, 365-366, 424-434, 456, 543, 618, 624, 695, 865, 867-869, 873-874, 876, 1024, 1152, 1196, 1198, 1201, 1203, 1205, 1207, 1210-1211, 1241, 1358, 1373, 1405, 1410, 1417, 1422, 1452, 1454, 1457-1458, 1460-1461, 1485-1489, 1492, 1498-1499, 1704-1705, 1710-1711, 1713, 1715, 1815, 1859, 1864-1866, 1871, 2021, 2024, 2049, 2051, 2069-2070
Interoperability Resources 425, 428, 430, 1486, 2051, 2070
Intracellular Current 249, 255, 257-260, 262-265
Introns 506, 511, 1168, 1172, 1178-1179, 1185, 1188, 1192
Iot 271, 975, 1366, 1401-1402, 1405-1407, 1409-1410, 1839, 1842-1847, 1849

K

KEGG 227, 557, 559, 1170
Ketones 13-14, 20, 22, 25, 199
Keystroke Dynamics-Based Authentication (KDA) 1260, 1282
KGML 557, 559
Kinaesthetic 1283-1288, 1290-1293
Kinect 726-731, 736, 738-739, 1006-1010, 1013-1015, 1017-1021, 1601-1604, 1609, 1617, 1966
K-Means 1306, 1310, 1312-1313, 1316-1318, 1322, 1325
Knowledge Creation 2, 150-151, 154
Knowledge Discovery 65, 169-170, 412, 421, 512, 979, 1025, 1544, 1558, 1742, 1994
Knowledge Factor 1282
Knowledge Management 150-151, 154-157, 162-165, 340-344, 349-351, 473, 670, 1816, 2070
Knowledge Management Process 150-151, 157, 162-165, 344, 350
Knowledge-based Systems 393, 395

L

Laboratory Information System (LIS) 1942
Laboratory Tests 55, 136, 172, 398, 411-412, 616, 1302, 1493, 1949-1950, 1991, 2001
Lancets 22
laparoscopic surgery 2036-2039
Laparoscopy 2038, 2047
Lattice Model 388
Layered architecture 460, 478, 480-481
Lethal Attacks 1799, 1812
Ligand 197-198, 203-205, 207-208, 520, 559, 1557-1561, 1564-1567, 1577
Limb Rehabilitation 1006-1007, 1010, 1018-1020
Living Labs 1980
locomotion classification 763-764, 766
Lysis 929

M

Machine Learning 31, 55, 202, 206, 326, 328, 335, 337, 391, 394-395, 410-412, 416-417, 420, 488, 503, 508, 510, 512, 523, 652, 708, 729, 769, 774, 1310, 1502, 1504-1505, 1508-1509, 1517, 1520-1521, 1826-1827, 1831-1833, 1845-1846, 1849, 1898, 1965, 1978-1979, 1989-1990, 1993-1995, 2001
Machine Learning Algorithms 335, 411-412, 512, 708, 1505, 1827, 1831, 1978, 1990, 1993-1994
Markov Chain 237-240, 242, 388
Mavisto 558
Mechanized Frameworks 1401
Medical Conjunctive Fuzzy Query 268, 275, 277
Medical Device 288, 290, 296, 853, 1343-1344, 1358
Medical Diagnosis 473, 518-519, 646, 682, 696, 1040, 1816, 1990, 1994-1995, 2001
Medical Information 80-81, 169-170, 185, 269-273, 361, 393, 473, 534, 665, 675, 699, 976, 1223, 1226, 1228-1229, 1232, 1234-1235, 1255, 1258, 1418, 1451-1453, 1457, 1460, 1468, 1477, 1674-1678, 1681, 1782, 1816, 1862, 1948, 1954, 2005, 2007
Medical Informed Consent 1223-1224, 1253
Medical Natural Language Processing 1504, 1521
Medication 15, 21, 23-25, 31-32, 34, 36, 59, 361, 395, 430, 487-488, 492-493, 495, 529, 615-616, 621-622, 624, 867, 874, 880, 886, 890, 1200, 1230, 1264, 1302, 1402, 1405, 1417-1418, 1452, 1504, 1506, 1745-1746, 1785, 1787-1788, 1795, 1807, 1847, 1850, 1946, 1949-1950, 1978, 1991-1992, 2057
Medicine 57, 59, 61-63, 67-68, 74, 76-77, 82-84, 86, 89, 94, 165, 197, 211, 234, 326, 349, 391-392, 394-395, 397-399, 425, 473, 475, 495, 532, 617-618, 677, 693, 844, 852, 866-867, 987, 989-997, 1001-1002, 1005-1006, 1028, 1155, 1159, 1162, 1200, 1223-1226, 1228, 1235, 1237, 1253, 1286, 1364-1374, 1420, 1504, 1766, 1816-1817, 1828,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

1847, 1861-1862, 1871, 1875, 1922-1923, 1942, 2003, 2019, 2033, 2049
MEMS 2008, 2041
Metagenomic 1168, 1170, 1172-1173, 1192
Metagenomic Gene 1168, 1192
Meta-heuristics 565, 568, 572
Mhealth 528-530, 533, 543, 546, 1404, 1410, 1603, 1619, 1922-1923, 1928, 1930-1931, 1933-1934, 1942
Microarray Dataset 646, 655, 707-708, 711, 713-715, 717
Microarray Medical Data 779, 781, 783, 809
Microarrays 227, 507, 519, 552, 554, 557, 581, 583, 1053-1054, 1058, 1063-1064, 1066, 1071, 1875, 1877
MICS Band 853, 860
miniature robots 2038
minimally invasive surgery 2034, 2036-2037, 2039-2040
Mobile Applications 86, 120, 437, 630-631, 881, 891, 1255-1256, 1258-1259, 1266-1270, 1272-1274, 1523, 1530, 1532, 1551, 1923, 1925-1927, 1951
Mobile Authentication 1256, 1259, 1262, 1268, 1270, 1272, 1275
Mobile Computing 86, 1254-1256, 1259, 1267, 1269-1270, 1272, 1275, 1282, 1451, 1633, 1922, 1934
Mobile Devices 86-90, 268-270, 284, 397, 437, 529, 616, 625, 665-666, 671, 1255-1257, 1261, 1269, 1272-1273, 1275, 1418, 1607, 1636, 1798, 1859, 1921-1931, 1942, 1944
Mobile Health (Mhealth) 270-271, 437-438, 451, 528, 665, 892, 1404, 1603, 1632-1635, 1641, 1644, 1647-1648, 1843, 1922, 1931, 1942
Mobile Health Services 1632, 1635, 1641, 1644
Mobile Healthcare 6, 271, 437-439, 442, 451-452, 609, 1254, 1256, 1263-1264, 1266, 1275, 1635
Model approximations 231
Modeling of Biological Systems 1919-1920
Modified Cat Swarm Optimization 781, 786-787
Modular Microrobots 2047
Molecular Computation 1919
Molecular Computing 1896, 1900, 1906
Molecular Operating Environment 1565-1566
Mrna 227, 235, 379, 388, 504, 507-508, 553, 1053, 1063, 1065, 1169, 1172, 1307, 1880-1881
Multi-Agent based Simulation 369-370
Multi-Agent System 874-875, 901, 929, 1025-1026
Multi-Annotator Gold Standard Agreement 815, 823, 825, 827-828
Multicellular System 371, 375, 378-379, 1903
Multi-Factor Authentication (MFA) 1273, 1282
Multi-objective optimization 514
Multi-Protocol Attacks 437, 453
Muscle Model 836, 838-839
Mutual Information 769, 1053-1054, 1056-1060, 1066, 1070-1071

N

Named Entity Recognition (NER) 326, 1503-1504, 1507, 1521
nanocomposite 1076
nanorobots 2041
Natural Computing 1920
Natural Language Processing (NLP) 326, 395, 1506, 1521
Natural Orifice Transluminal Endoscopic Surgery 2039, 2047
Nearest neighbor method 50, 55
Nerve Fiber Stimulation 249, 254, 257, 265
Network Motifs 551, 555-556, 558, 564, 1908
Neural Networks 50, 58, 202, 337, 391-395, 397, 421, 472, 487-488, 503, 511, 519, 651, 653, 655, 676, 678, 686, 781, 789, 836, 1312, 1328, 1833, 1836, 1990, 1994
Neural-net Ensemble 662
Neuroimaging 1963-1965, 1972-1976
Neuromuscular Rehabilitation 837, 845
Next Generation Sequencing 581, 1168, 1171, 1173, 1181, 1192, 1995
NICU 288-291, 293-294, 296
Nonadherence 487-488, 491-492, 494-497
NOTES 7, 62, 170, 172, 325-326, 345, 348, 393, 395, 397, 411, 1199, 1204, 1286-1287, 1302, 1387, 1418, 1420, 1506, 1686, 1727-1728, 1742, 1760, 1848, 1928, 1991-1992, 2036, 2039
Nursing Practice 865, 873-875
Nursing Student 1678, 1681
NVIDIA Jetson TK1 629-631, 633-639

O

Observability 1426-1428, 1449
Ontology Evolution 1024-1026, 1031-1033
Ontology Learning 1024-1025
Open Data 1157-1158, 1160, 1162, 1167, 1532, 1966, 2069
Open Source 83, 203, 363, 556, 598, 602, 716, 829, 904, 1236, 1242, 1391, 1396, 1513, 1710, 1831-1834, 1836, 1932, 1996, 2057
Oral Glucose Tolerance Test 16-17, 40
Ordinary Differential Equations 244, 388, 899

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

Organization Performance 1583
OWL 173-175, 475, 478, 529, 534-535, 537, 540, 542, 1026, 1031, 1242, 1817, 2050, 2052-2053, 2055

P

Pancreas 13-15, 17-20
Parallel Computation 1832
Parallel Computing 1065, 1067, 1560, 1920
Parallel Processing 93
Paralysis 833, 835-837
Participating Research Consent 1228-1229, 1253
Particle Swarm 502, 513-514, 517-518, 565, 568-570, 728, 743-745, 749-753, 758, 780, 783, 785, 1313-1315, 1318, 1327, 1906
Particle Swarm Optimization (PSO) 502, 513-514, 518, 568, 758, 1318, 1906
Pathology 172, 293, 672, 1157, 1387, 1419, 1876, 1922-1926, 1928, 1934, 2038
Patient Centered Medical Home 1946-1947
Patient Monitoring 85, 270, 456-460, 466, 665, 931, 938, 1633, 1965, 1970, 2003, 2015
Patient Protection and Affordable Care Act 1859, 1943-1944
Patient Safety 2, 151, 174, 496, 1223-1224, 1299, 1418, 1451, 1700, 1739-1741, 1749, 1761, 2019, 2034
PCA 202, 517, 582, 584-585, 589, 711, 713, 716-718, 722, 772, 780, 1055-1057, 1061-1062, 1066, 1071, 1312, 1587
PCEH 136, 149
Performance Dashboards 1745
Perioperative Process 1738-1740, 1744-1745, 1753
Personal Health Records 135, 528, 598, 610, 1255-1256, 1258, 1452, 1787
Personalized Medicine 211, 1235, 1364-1374, 1862, 1871, 1875
Pervasive computing 268-270, 456-457
Pervasive Healthcare 268-272, 457, 460, 609
PGOT Framework 1764-1765, 1774, 1776
Pharmaceutical 195, 487-488, 491, 496, 645, 1369, 1410, 1506, 1532, 1561, 1842, 1844, 1849, 1991
Phase Locking Value (PLV) 1654, 1656, 1659
Photobacterium Leiognathi 908, 910, 928-929
Physical Activity 14, 18, 32, 40, 1599-1602, 1605, 1607, 1609-1613, 1618-1619, 1969
Physical Healthcare Crimes 1812
physical training 1599, 1607, 1609-1610, 1618-1620, 1977
Plasma processing 1082
Play-Doh 1283, 1285, 1287, 1290-1293
Pluralistic Healing Method 987, 1005
Polycentric Governance 1708, 1713-1716
Population Health Management 74-76, 84-85, 90, 93-94
Possession Factor 1282
Postmortem Analysis 1794, 1812
Postnatal Depression 879-880, 884, 891
Postnatal Depression Intervention and Support 879
Power Spectrum Densities (PSD) 1655
predefined messages 1613-1615
Predictive Modeling 410-412, 414, 421, 1993
Prescription 23, 39, 76, 82, 119, 123, 487-488, 491, 494, 496, 615, 621-622, 624, 1265, 1269, 1300, 1365, 1387, 1406, 1420, 1451, 1454, 1493, 1760, 1764, 1782, 1788, 1792, 1828, 1843, 1848, 1980, 1991-1992, 1995
Primary Care 50, 59, 75, 1265, 1297-1302, 1455, 1713, 1764, 1766, 1787, 1946-1948, 1964-1965
Principal Component Analysis 128, 202, 582, 584, 711, 713, 725, 772, 780, 1013, 1055, 1312, 1587, 1827
Privacy Consent 1226-1227, 1231-1232, 1240, 1242, 1246, 1253
Probabilistic Boolean Networks 226, 228, 235, 244
Process Framework 150-151, 157-158, 161-165
Process Mining 424-427, 431-432, 434
Progressive alignment 567-568, 632
Protected Health Information (PHI) 51, 64, 183, 1342, 1502-1503, 1949-1950
Protein Motif 507, 1312, 1322, 1327
Protein Sequence 505-506, 509, 565, 569, 691, 1306-1307, 1311-1314, 1316, 1318-1322, 1325, 1327, 1558, 1904
Protein-Pollutant Interaction 1558
Proteomics 581, 583-584
PSO 502, 513-520, 525, 527, 568-572, 749, 752, 756-758, 781, 1306, 1312-1314, 1316-1319, 1322, 1325, 1327, 1906
Pulsed laser deposition (PLD) 1077, 1087

Q

QD 1878, 1894
Qos 87, 596, 611, 973, 1403-1404, 1407, 1787
Qualitative Study 1585, 1611, 1708, 1732
Quality Control 580-582, 585, 1062, 1480, 1877, 1921, 1925, 1931, 1942, 1991

R

Random Forest 202, 410, 412, 416-417, 709, 780, 1009
Random Projection 709, 713, 725
RDF 104, 172-173, 474-476, 478, 528-529, 533-535, 537, 543, 545-546, 549, 1817, 1846, 1966, 2050-

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

2056, 2058, 2065
Reaction Diffusion Systems 389
Real Time Analysis 1839, 2007
Reduced order 228
Regional ICT Governance 1708, 1710, 1712-1715
Regression 50, 55, 58, 202, 410, 412, 416-417, 465, 781, 789, 1054, 1058, 1060, 1062, 1066, 1068, 1070-1071, 1416, 1428, 1435-1437, 1441, 1444-1445, 1832, 1994
Relative Advantage 1416, 1425-1428, 1435-1436, 1440-1445, 1449
Resistance Traditionalism 1001
Reverse Engineering 427, 1053-1054, 1058, 1060, 1066, 1068, 1070-1071
RF magnetron sputtering 1077
Risk-Based Authentication (RBA) 1282
Ritual Purification 987, 991, 1005
RNA 226, 504, 507-509, 511, 552, 566, 1057, 1062, 1064, 1169, 1171-1172, 1192, 1876, 1878-1883, 1894, 1904, 1919
Robotic Surgery 2047
RT-PCR 1875, 1877, 1894
Rule-Based 393, 424, 430, 1026, 1504-1505, 1508-1509

S

SBML 557, 559, 561-562
Scenarios Reconstruction 1808
Science of Individuality 74, 76-77, 88, 90, 94
Security Level 1487-1488, 1496, 1499
Security Policies 185, 1260, 1498, 1684, 1686-1687, 1692, 1694, 1697-1699, 1701-1702
Semantic Interoperability 169, 171-172, 424, 429, 431-434, 543, 1458, 1461, 1705, 2051
Semantic Matching 471-473, 476, 1814, 1816-1819
Semantic MediaWiki 2048-2049, 2052, 2057-2058, 2069-2070
Semantic Ontology Store 176-179
Semantic Role Labeling 1023-1024
Semantic Web 107, 170, 172-173, 471, 528, 1023, 1209, 1814-1815, 1817, 2048-2057, 2059, 2068-2071
sensor fusion 764, 774
Sentiment Analysis 104-105, 107, 109, 112, 1024, 1833
Sequence Alignment 509-510, 520, 565-569, 571, 629, 631-632, 639, 900, 1170, 1181
Serious Games 1602, 1607, 1963, 1965-1966, 1969
Service Quality 151, 342, 664, 1195, 1370, 1374, 1460, 1582, 1586, 1588, 1590-1591, 1632-1637, 1644, 1860, 1943, 1945-1947, 1952, 1955
Sharps Bucket 22-23
Signal Processing 294, 666, 688, 690, 748, 762-766, 768-769, 774, 816, 818, 820, 931-932, 937, 939, 963-964, 1654-1655, 1671, 2008, 2010
Signed Graph 1173-1176, 1192
Significant Amino Acids (SAA) 1325, 1327
Single Board Computers 598, 601, 610
Smart City 1403, 1407
SOA 133-140, 142, 144-145, 149, 458, 624-625
SOA Governance 133-140, 142-145, 149
Social Media 882, 1200, 1537, 1539-1543, 1547, 1551, 1648, 1758-1762, 1764, 1768, 1771-1772, 1774-1777, 1827
Spatial Analysis 1524, 1532
Speech Imagery 1654-1659, 1662-1664, 1667, 1669-1671
Sphygmogram 666, 706
Spreading Activation 471-473, 476, 480-481, 1814-1816, 1819-1820, 1822
Statistical Moments 1328, 1330-1331, 1333-1334
STEM 331, 371, 376, 397, 1054, 1083, 1283-1285, 1290, 1292-1293
Stochastic Master Equations 227
String Matching 298-299, 319, 321, 2055-2056
Stroke Rehabilitation 837, 840
Sub-Graphs 555, 558, 564
Success of ICT Projects 2023
Support Vector Machine (SVM) 50, 55, 208, 290, 292, 332, 394-395, 410, 412, 510, 682, 684, 727, 729, 766, 773, 779-780, 795, 1312, 1329, 1505, 1508, 1521, 1994
surgical robots 2035, 2038, 2041
Surgical Site Infection 410-411
Swarm Intelligence (SI) 502, 524
Symmetrical Uncertainty 711
Synonymy 474, 528, 537, 545-546
Synthetic Biology 1878-1880, 1883, 1908-1910
System Design 1-2, 4-7, 9, 609, 1018, 1301, 1303, 1613
System Security 1227
Systems Biology 369-375, 377, 382, 384, 388, 561-562, 1899, 1906, 1910, 1919-1920

T

Telecytology 1921, 1925, 1927, 1934, 1942
Telediagnosis 1404, 1925, 1942
Term Extraction 1024, 1026
Text Mining 63, 68, 107, 1521, 1995
Thermography 1329, 1331
Top-Down Approach 427, 434, 1058, 1068, 1707, 1865
Top-K Answers 268
Top-N Rewritings Algorithm 268-269, 273, 279-280,

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use

282-285
Traditional Medicine 992-995, 997, 1005
Traditional People 992, 995-998, 1002, 1005
Transcranial Magnetic Stimulation 249-250
Treatment Adherence 879, 881-884, 891
Treatment Consent 1224-1227, 1230-1231, 1240-1241, 1246, 1253
Trialability 1426, 1428, 1449
Trusted Platform Module (TPM) 437, 452

U

Universal Integrated Circuit Card (UICC) 452
Use Of Information Systems 1203, 1684, 1720
User Authentication 1254-1260, 1262, 1266-1270, 1272-1273, 1275, 1282, 1406, 1458, 2012
User Profiling 1256, 1259, 1282
UTAUT 1633, 1635-1636, 1642, 1722-1723

V

Valuesets 432, 2048, 2051, 2056-2057
Virtual Screening 194-195, 201, 203
Virtual Slide 1931-1932, 1942

W

Waiting Time 89, 284-285, 616, 890, 1060, 1522-1524, 1527-1528, 1530-1532
Wearable 89, 456-457, 459, 461, 529, 595, 667-668, 672, 706, 727, 745-747, 751, 758, 844, 971, 973-976, 1007, 1255, 1258, 1364, 1369, 1403-1405, 1407, 1412, 1607, 1789, 1808, 1829, 1841, 1843-1847, 1864, 1871, 1992, 2008
Well-Being 76, 90-93, 139, 292, 357, 457, 488, 876, 1601, 1607, 1609, 1870, 1944, 1948, 1954, 1978
Wiki 74, 94, 2048, 2052, 2057, 2059, 2066-2067, 2070-2071
Wikitext 2052, 2057-2058, 2063, 2066-2068, 2071
Wireless Body Area Networks 459, 971, 973, 1782
Wireless Body Sensor Networks 31, 664, 706, 981, 1404, 1406
Wireless Sensor Networks (WSN) 462, 744, 980, 1782, 1785, 2002, 2010
Wireless Sensor Networks for Healthcare 1785
Workflow Management 196, 1225, 1231, 1235-1236, 1242-1243, 1246, 1253

X

Xefteris 1968, 1970

Z

Zigbee (802.15.4) 748

EBSCOhost - printed on 2/10/2023 8:17 PM via . All use subject to https://www.ebsco.com/terms-of-use