# TECHNOLOGY AND INFORMATICS 221
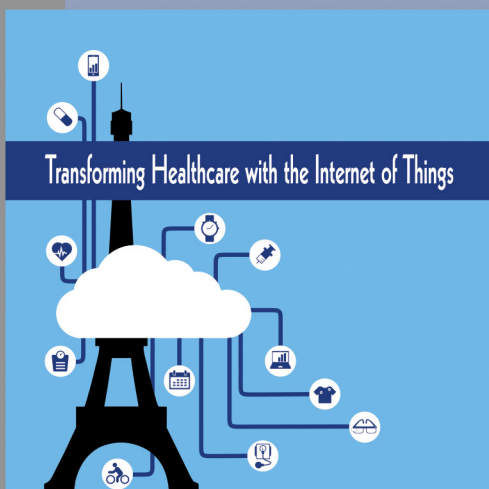
# Transforming Healthcare with the Internet of Things

*Proceedings of the EFMI Special Topic Conference 2016*



Transforming Healthcare with the Internet of Things

Editors: Jacob Hofdijk
Brigitte Séroussi
Christian Lovis
Frédéric Ehrler
Floor Sieverink
Adrien Ugon
Mira Hercigonja-Szekeres

*IOS Press*

# TRANSFORMING HEALTHCARE WITH THE INTERNET OF THINGS

# Studies in Health Technology and Informatics

This book series was started in 1990 to promote research conducted under the auspices of the EC programmes' Advanced Informatics in Medicine (AIM) and Biomedical and Health Research (BHR) bioengineering branch. A driving aspect of international health informatics is that telecommunication technology, rehabilitative technology, intelligent home technology and many other components are moving together and form one integrated world of information and communication media. The series has been accepted by MEDLINE/PubMed, SciVerse Scopus, EMCare, Book Citation Index – Science and Thomson Reuters' Conference Proceedings Citation Index.

## Volume 221

*Recently published in this series*

# Transforming Healthcare with the Internet of Things

Proceedings of the EFMI Special Topic Conference 2016

Edited by

## Jacob Hofdijk

*Casemix – CQT, Village Foundation, Utrecht, The Netherlands*

## Brigitte Séroussi

*Sorbonne Universités, INSERM, Paris, France*

## Christian Lovis

*Division of Medical Information Science, University Hospitals of Geneva, Switzerland*

## Frédéric Ehrler

*Division of Medical Information Science, University Hospitals of Geneva, Switzerland*

## Floor Sieverink

*Centre for eHealth and Wellbeing Research, Department of Psychology, Health & Technology, University of Twente, Enschede, The Netherlands*

## Adrien Ugon

*Sorbonne Universités, INSERM, Paris, France*

and

## Mira Hercigonja-Szekeres

*University of Applied Health Studies, Zagreb, Croatia*

## IOS
### P r e s s

Amsterdam • Berlin • Washington, DC

LEGAL NOTICE

The publisher is not responsible for the use which might be made of the following information.

PRINTED IN THE NETHERLANDS

# Preface

This volume contains the proceedings of the Special Topic Conference (STC) of the European Federation for Medical Informatics (EFMI). The organisation of the STC is part of a long tradition of EFMI working groups to organise scientific events focused on important trends in medical informatics and eHealth. In 2016, the special topic is "*Transforming Healthcare with the Internet of Things*" in relation to the EFMI working group Personal Portable Devices (PPD). STC 2016 takes place in Paris, France, organised by the *Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé* (LIMICS) under the auspices of EFMI and the French Association for Medical Informatics (AIM).

Only a few years ago, devices were limited to health cards and personal portable devices. Since then, devices have dramatically evolved to include wearables, sensors, and actuators for measuring health values. The application of such technologies in the field of health, social care and wellness has attracted the attention of both patients and members of the general public interested in supporting or improving their health and wellbeing. One of the characteristics of these 'devices' (sometimes too small to observe thanks to nanotechnology) is to be 'connected' and to communicate with other connected devices and systems. This has been the game changer, as it replaces the cumbersome and often error-prone intervention of the human being who was previously necessary to enter data.

The Internet of Things (IoT) is thus turning out to have a major impact on the information paradigm in healthcare. The patient can now become their own Chief Operational Officer, as described by Eric Topol in his recent book *The Patient will see you now*[1]. By providing tools that are able to generate large quantities of data that must be processed in real time, the IoT can have a potentially transformative effect on healthcare, allowing medicine and patient management to evolve from a discrete encounter-based process to a continuous patient-empowering management. This requires an adequate answer from traditionally organised healthcare, which will have to find ways to address new challenges related to respect for consumer privacy, cyber security and data integrity.

STC 2016 deals with the convergence of a process originally fuelled by technical and scientific forces, and the current political forces driven by the sustainability agenda of health and social care. This emphasises the process to humanise the individual who is more and more connected and surrounded by the IoT. Our ambition is to concentrate on this debate and create a platform for these different dimensions of this unstoppable development.

As a conclusion, we quote the Blue Line Statement presented at the 31th PCSI Conference in The Hague on October 16th 2015 in order to emphasise the transition of health and social care systems and the shift from care to citizen-driven health:

> *In order to achieve meaningful improvements in the health of the population, it is essential to understand the combination of health and social care issues for people. This requires health and social care systems to be interoperable, both from*

---

[1] Eric Topol, The Patient Will See You Now, Basic Books, New York, 2015.

*a technological and semantic point of view. Care should be aligned around the person and strive for social interoperability between the professions serving them, and systems must be designed with empathy and respect as core underpinning values. The pillars of the Blue Line model represent key principles for the design and delivery of person-centred, integrated care systems. We encourage policymakers and health system leaders to adopt these principles and create the societal incentive framework to enable this vision to be realised. The Village Track participants of the PCSI Conference recommend that further support be sought to continuously develop and formalise the Blue Line principles as requirements for supporting holistic, person-centred, integrated care systems in the Netherlands, and beyond.*

## STC Programme

The call for papers has resulted in 70 submissions from 30 countries, which were peer-reviewed by over 160 highly appreciated experts of the EFMI biomedical informatics network. Over 400 authors and co-authors are involved in the accepted contributions that shape the programme. The conference starts with a key contribution from Bernard Benhamou entitled Internet of Things & Medicine: A European Perspective. Then, Peter Pharrow explains how 'We are entering the era of the Internet of (every)-Thing'. Christian Lovis, as the third keynote speaker, ends the conference by addressing the key to 'Moving from a care-driven system to a health-centred paradigm: active objects, data, and information'.

The reviewing process has demonstrated that, up to now, the scientific work related to the IoT in healthcare has been focused on a mix of technologically-driven issues, but the potential to reform the health and social care systems is still underexploited. This can be seen as a positive sign, as the scientific world first wants to understand and prove the potential of the IoT before widely implementing it within the healthcare system. Accepted contributions reflect the scientific work on the impact of the IoT and the societal dimensions of the IoT in the sessions related to the transformation of healthcare, while the sessions on ontologies, decision support, clinical information systems, and data reuse complete the programme. To enrich the scientific sessions, the conference also offers tutorials, workshops, posters, short communications, demonstrations and a plenary panel.

From Sunday 17th to Tuesday 19th April 2016, Paris will not only be the city of light, but also the centre for world-citizens involved in the changing of traditional health care. It is with respect for the history of French medicine that the LIMICS chose the "*Ecole de médecine*" (created in 1794), as the historic, almost sacred, venue for STC 2016 to discuss the future of healthcare.

On behalf of the Scientific Program Committee

Jacob Hofdijk

SPC team
Brigitte Séroussi, Christian Lovis, Floor Sieverink, Frédéric Ehrler and Adrien Ugon

# List of Reviewers

Jos Aarts
Dionisio Acosta
Leila Ahmadian
Elske Ammenwerth
Yesim Aydın Son
Lejla Begic Fazlic
Johan Gustav Bellika
Mohamed Ben Said
Lino Bessonart
Isabelle Bichindaritz
Jürg Blaser
Damian Borbolla
Alessio Bottrighi
Niels Boye
Ruth Breu
Sandra Bringay
Ramón Camano-Puig
Werner Ceusters
Jean Charlet
Carlo Combi
Ronald Cornet
Catherine Craven
Manfred Criegee-Rieck
Vasa Curcin
Margaret Czart
Tina Dannemann Purnat
Berry de Bruijn
Nicolette De Keizer
Lucio Tommaso De Paolis
Kerstin Denecke
Thomas M. Deserno
Michel Dojat
Paul Donaldson
Catherine Duclos
Martin Dugas
Claudio Eccher
Frédéric Ehrler
Eric Eisenstein
Amado Espinosa
Vahid Farahmandian
Jesualdo Tomás Fernández-Breis
John Finnell
Mircea Focsa

Nils Daniel Forkert
Fleur Fritz
Elia Gabarron
Matthias Ganzinger
Gersende Georg
Andrew Georgiou
Bernard Gibaud
Francesco Giuliani
Anita Ground
Wided Guedria
Werner Hackl
Vigdis Heimly
Harald Heinzl
Kai Heitmann
Jacob Hofdijk
Vojtech Huser
Nico Jähne-Raden
Markus Jochem
Martti Juhola
Nikitas Karanikolas
Thomas Karopka
Finn Kensing
Josipa Kern
Soudabeh Khodambashi
George Kim
Won Kim
Petra Knaup
Peter Kokol
Vassiliki Koufi
Vassilis Koutkias
Lill Kristiansen
Hallvard Lærum
Frederique Laforest
Antti Lahtela
Luis Lapao
Thierry Lecroq
Giorgio Leonardi
Lenka Lhotska
Siaw-Teng Liaw
Christian Lovis
Diana Lungeanu
Martin Lysser
Farah Magrabi

Brad Malin
John Mantas
Romaric Marcilly
Miguel Angel Mayer
Oleg Mayorov
Luca Mazzola
George Mihalas
Martina Mittlboeck
Dr. Mohyuddin
Alberto Moreno-Conde
Fleur Mougin
Henning Müller
Maurice Mulvenna
Peter Murray
Eustache Muteba Ayumba
Sahiti Myneni
Hirenkumar Nakawala
Øystein Nytrø
Frank Oemig
Andrej Orel
Wolfgang Orthuber
Mustafa Ozkaynak
Alessia Paglialonga
Paulo Paiva
Carlos Luis Parra Calderón
David Parry
Sylvia Pelayo
Petra Perner
Francesco Pinciroli
Monika Pobiruchin
Mihail Popescu
François Portet
Luigi Portinale
Michael Prinz
Thomas Rindflesch
Christoph Rinner
Jean Marie Rodrigues
Alejandro Rodriguez
Samuel Rosenbloom
Ann Kristin Rotegard
Etienne Saliez

Raquel Santos
Antony Sara
Kaija Saranto
Gerd Schneider
Björn Schreiweis
Philip Scott
Martin Sedlmayr
Danielle Sent
Brigitte Séroussi
Michael A. Shifrin
Chris Showell
Floor Sieverink
Dean Sittig
Neil Smalheiser
Lina F. Soualmia
Stéphane Spahni
Tom Sparrow
Peter Spyns
Martin Staemmler
Holger Stenzhorn
Lăcrămioara Stoicu-Tivadar
Vasile Stoicu-Tivadar
Sanders Su
Selma Supek
György Surján
Özgür Tosun
Shuichi Toyoda
Guenter Tusch
Adrien Ugon
Erik van Mulligen
Jan Vejvalka
Vivian Vimarlund
Dongwen Wang
Patrick Weber
Martin Wiesner
Alfred Winter
Melissa Yorks
Dimitris Zikos
Blaz Zupanand
Jana Zvarova

# Contents

## Ontology and Decision Support

## Clinical Information Systems and Data Reuse

## Posters

# Transforming Healthcare with the IoT

This page intentionally left blank

*Transforming Healthcare with the Internet of Things*
*J. Hofdijk et al. (Eds.)*

3

# Using Gamification Combined with Indoor Location to Improve Nurses' Hand Hygiene Compliance in an ICU Ward

Luís Velez LAPÃO[a,1], Rita MARQUES[a,b], João GREGÓRIO[a], Fernando PINHEIRO[c], Pedro PÓVOA[c], and Miguel MIRA DA SILVA[b]

[a] *Global Health and Tropical Medicin, Instituto de Higiene e Medicina Tropical – Universidade Nova de Lisboa, Lisboa, Portugal*
[b] *Instituto Superior Técnico – Universidade de Lisboa, Lisboa, Portugal*
[c] *Centro Hospitalar de Lisboa Ocidental – Hospital S. Francisco Xavier, Lisboa, Portugal*

**Abstract.** Healthcare acquired infections are among the biggest unsolved problems in healthcare, implying an increasing number of deaths, extra-days of hospital stay and hospital costs. Performing hand hygiene is a simple and inexpensive prevention measure, but healthcare workers compliance with it is still far from optimal. Recognized hurdles are lack of time, forgetfulness, wrong technique and lack of motivation. This study aims at exploring gamification to promote nurses' HH compliance self-awareness and action. Real-time data collected from an indoor location system will provide feedback information to a group of nurses working in an ICU ward. In this paper both the research's motivation and methods is presented, along with the first round of results and its discussion.

**Keywords.** Internet of Things, Healthcare acquired infections, indoor location, gamification, nursing, design science research

## Introduction

Healthcare acquired infections (HAI) are infections that are neither present nor incubating when a patient is admitted to hospital [1][2]. HAI are a risk that hospitals must control to manage healthcare economically and safely for patients who can become disabled in the long-term or even die. Although preventable, by means of hand hygiene (HH) compliance, these infections are the most frequent adverse event a patient can experience during care delivery, and cause more deaths than AIDS, breast cancer and car accidents together [3]. Nonetheless, leading busy healthcare workers (HCW) to comply with HH remains puzzling. Recognized hurdles are lack of time, forgetfulness, wrong technique and lack of motivation. Besides, nurses' perception about their compliance is often disturbed by a busy schedule. Therefore, it becomes crucial to monitor nurses' compliance with existing guidelines and provide them with feedback regarding their performance. Direct observation, the observation of HCW's

---

[1] Corresponding Author.

HH practice by professional observers, is the standard approach to fulfil this task, but it is costly and time-consuming. Hospitals need to come up with innovative ways of doing this. Automated monitoring systems have emerged during the last few years, and can electronically identify when an HCW uses a sink or a hand rub dispenser. It provides exact quantitative results, which can be used to examine trends regarding the value of HH compliance over time.

Gamification is a recent but popular approach which can be defined as "the use of game elements and game-design in non-game contexts" [4] to "engage and motivate people to achieve their goals" [5], providing a whole different user experience. It aims at stimulating people's intrinsic motivation in doing an activity by making it rewarding for itself. Game elements are the "toolkit" for building a game [4], they must be chosen at the end of the game design process, after some variables are analyzed and defined (goals, behaviors we want to stimulate, target players, etc.). Wabash and Hunter provide a list of game elements divided into three categories with different levels of abstraction [4].

This study aims at exploring the use of gamification to promote nurses' HH compliance self-awareness and action. An automated monitoring system is be used to collect data in real time and provide feedback information to a group of nurses working in an ICU ward, in a fun and engaging way.

## 1. Methods

The solution, which consists on an automated monitoring system composed by an indoor location system and a gamified system is iteratively designed, tested and evaluated using a design science research approach [6]. The solution was presented to its target users (the nurses) and its usage was simulated in both artificial setting (by voluntary users, i.e. the researchers) and clinical environment (by nurses).

### 1.1. Design and implementation of the solution

The first component of the solution is an innovative indoor system based on smart beacon's technology. This was developed together with a Portuguese startup, focused on the development of solutions related with "Internet of Things" that communicates using Bluetooth and a proprietary protocol (operating on the 2.4GHz frequency band). The system was built using a proximity based technique. More specifically, the smart beacons receive information from smart tags (carried by nurses) and they send a message to the server (communicating its position, the smart tag detection and the current time) whenever a smart tag is detected approaching the beacon or walking away from it. Analyzing the messages stored in the server, we are able to detect nurses' positions over time.

The next step is for the system to detect and validate HH moments. To achieve this, we use the World Health Organization (WHO)'s "My five moments for hand hygiene" framework [7], which links specific moments to HH opportunities.

With this framework, we are able to create and implement business rules in our system (for example, if a nurse is approaching a bed, it must have approaching an alcohol hand rub dispenser or a sink previously). After this processing, we have information about each nurse's compliance, which is displayed, in an anonymous way, in a screen in real-time.

This, along with other elements, including the player's profile, composes the gamification solution which aims at solving the compliance problem by engaging and motivating people to achieve specific (and preferably pre-determined) goals [8], using several and distinct game elements (feedback, competition, points, levels, badges, etc.).

The technological architecture of this solution is presented in Figure 1.
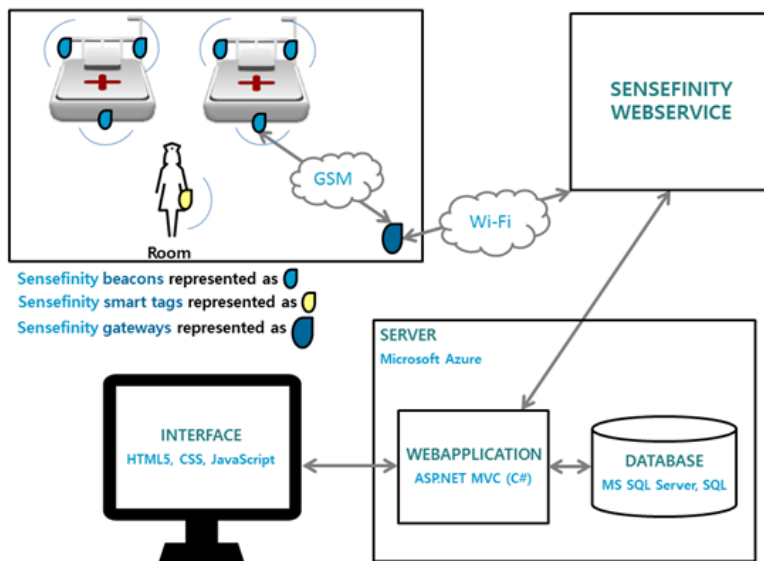


**Figure 1.** Technological architecture of the solution designed.

## 1.2  Meeting with the users

There have been some meetings with the nurses from the UCI. We presented the Information System (IS) aiming at gathering feedback regarding their feelings about it.

## 1.3  Simulation and tests

In order to validate if the solution was technically practical, it was tested in a simulated environment by voluntary users composed of the research members. A protocol was written and executed and the computed HH compliance rate was compared with the expected rate. To analyze the impact of the IS's usage by our target user, four nurses from the ICU ward were asked to carry a smart tag during a workday. This data was compared to a previously established baseline (in respect to an observational study) to measure the behavioral changes.

## 2. Results

## 2.1  Installation of the system

After being designed and implemented, the IS has been installed in the ward. A screen was positioned in the nurses' room and 26 beacons fixed at specific positions: in the rooms' doors, in each alcohol-based hand rub container, in each sink and in each bed.

By doing so, the system was able to trace a nurse's position along time based on the proximity to each beacon.

## 2.2 Feedback from the nurses

The group of nurses to whom the IS was presented to test the concept reported that they considered this system as a unique opportunity to receive feedback about their performance. Although they are sometimes subject to audits, they reported that the system would give them a totally different experience. Even though worried by the accuracy of the location system, they found the avatars experience amusing. About future improvements, they showed little interest in components like badges, virtual goods and content unlocking because it would require them to use the system outside their labor hours. They, however, liked the concept of leaderboards.

To conclude, we asked them if they would prefer to maintain their privacy or if they would like their name (or a chosen nickname) to appear on the screen. They said that this was indifferent for them, since they had no problem in having their identity exposed on the screen.

## 2.3 Simulation

The simulation in a non-clinical setting by the researchers presented good results. The system worked accordingly to our expectations, returning a HH compliance rate of 100%. Regarding the simulation performed by the nurse from the ICU ward during one 12-hour shift, the feedback received was that she got happier as she progressed in the game, and whenever she noticed that the rate had decreased (even if only a little), she felt the urge for being more aware of the HH moments.

## 3. Discussion

Since we were able to detect the nurse's movements using proximity and to quantify the compliance with a very good precision, the IS was conceptually validated. This wouldn't be possible without the indoor location solution that provided the nurse's position with accuracy in a real-time basis.

The participating nurses recognized the experience as an opportunity to improve their performance, which corresponds with their expectations collected during the initial meeting. Since nurses were a little skeptical in using some game elements that required them to access the system outside their work time, we decided to include a functionality in the system to send an e-mail to each nurse at the end of the day. This simple e-mail provides feedback regarding their HH compliance rate and has a link to their profile for further information. One of the nurses realized that she ended up consulting the webpage, even though she said she wouldn't in a first instance. The impact of gamification on HH compliance is still under evaluation. Even though we only performed low scale validation tests to check whether or not the concept would work, so far the results show that the IS is a promising solution in improving nurses' awareness.

A demonstration in the ICU ward is already planned. During a 5-days trial, 24 nurses will be using the gamification solution and will be provided with feedback regarding their HH compliance rate. Simultaneously, we will observe their behaviors and reactions, trying to understand if they are comfortable with the system and if they trust the presented results. We will also focus on spotting technical issues that might be

leading to undesired side effects of the system. In the end, we will analyze the gathered results and refine our tool, both in terms of improving our gamification solution and fixing some problems that may emerge during the demonstration. After this larger evaluation, our goal is to implement the solution in a unit of another hospital for a longer period.

To conclude, we believe that the IS is aligned with nurses' needs and that it has a positive impact on their daily routine. Although we noted some resistance to some ideas we discussed for future work with the nurses. The nurse who participated in our initial test agreed that she ended up using the system, so we trust that after using the IS on a daily basis nurses can be more interested in different options. It must be highlighted that the nurses participated from the beginning in the IS's design, which enabled a higher sense of ownership in the process, recognized as a factor to improve performance.

## Acknowledgments

## References

[1]  World Health Organization, "Who guidelines in hand hygiene in health care: a summary clean hands are safer hand," 2005.

[2]  J. Garner, W. Jarvis, T. Emori, T. Horan, and J. Hughes, "CDC definitions for nosocomial infections," APIC Infect. Control Appl. Epidemiol. Princ. Pract., vol. 16, no. 3, p. A–1–A–20, 1996.

[3]  C. Carboneau, E. Benge, M. T. Jaco, and M. Robinson, "A lean Six Sigma team increases hand hygiene compliance and reduces hospital-acquired MRSA infections by 51%.," J. Healthc. Qual., vol. 32, no. 4, pp. 61–70, 2010.

[4]  K. Werbach and D. Hunter, For the Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press, 2012.

[5]  B. Burke, Gamify: How Gamification motivates people to do extraordinary things. Bibliomotion, Inc., 2014.

[6]  K. Peffers, T. Tuunanen, M. a. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," J. Manag. Inf. Syst., vol. 24, no. 3, pp. 45–77, 2008.

[7]  World Health Organization, "Guidelines on hand hygiene in health care: first global patient safety challenge. Clean care is safer care," World Health Organization, 2009.

[8]  D. King, F. Greaves, C. Exeter, and A. Darzi, "'Gamification': Influencing health behaviours with games," *J. R. Soc. Med.*, vol. 106, no. 3, pp. 76–78, 2013.

# Integrated System for Monitoring and Prevention in Obstetrics-Gynaecology

Andreea ROBU[a,1], Bianca GAUCA[a], Mihaela CRISAN-VIDA[a], and
Lăcrămioara STOICU-TIVADAR[a]

[a]*Department of Automation and Applied Informatics, University Politechnica
Timisoara, Romania*

**Abstract.** A better monitoring of pregnant women, mainly during the third trimester of pregnancy and an easy communication between physician and patients are very important for the prevention and good health of baby and mother. The paper presents an integrated system as support for the Obstetrics – Gynaecology domain consisting in two modules: a mobile application, ObGynCare, dedicated to the pregnant women and a new component of the Obstetrics-Gynaecology Department Information System dedicated to the physicians for a better monitoring of the pregnant women. The mobile application informs the pregnant women about their status, permits them to introduce glycaemia and weight values and has as option pulse and blood pressure acquisition from a smart sensor and provides results in a graphic format. It also provides support for easy patient-doctor communication related to any health problems. ObGyn Care offers nutrition recommendations and gives the pregnant women the possibility to enter a social space of common interests using social networks (Facebook) to exchange useful and practical information. Data collected from patients and from sensor are stored on the cloud and the physician may access the information and analyse it. The extended module of the Obstetrics-Gynaecology Department Information System already developed supports the physicians to visualize weekly, monthly, or on a trimester, the patient data and to discuss with her through the chat module. The mobile application is in test by pregnant women and medical personnel.

**Keywords.** pregnancy, obstetrics-gynaecology department, android, glycaemia, blood pressure, monitoring, prevention

## Introduction

Creating a continuous process between patient, obstetrician and neonatal physicians, in the same time empowering the patient is important for a normal and healthy time during pregnancy [1]. Each part involved is as important in baby survival and healthcare status and may decrease neonatal mortality and morbidity through collecting critical data of the pregnant women. Pregnancy brings about many changes in a woman's body. A pregnant woman should have her weight monitored regularly during pregnancy, because every risk of pregnancy, both to the mother and to the baby, is increased with maternal obesity. Obesity in pregnant women can lead to all sorts of complications, including the death of the mother or of the baby through stillbirth, the baby having foetal abnormalities, the woman suffering pre-eclampsia or gestational diabetes. Around 35 per cent of women who die in childbirth are obese, while 30

---

[1] Corresponding Author.

percent of pregnant women are overweight or obese [2]. Measuring the blood pressure is also very important, representing a way of telling how well your pregnancy is going, watching for signs of a potentially serious complication called pre-eclampsia, particularly later on in your pregnancy. Pre-eclampsia isn't fully understood, but it's thought it to happen when the placenta is not working as well as it should. This can lead to high blood pressure and other problems [3]. Another risk factor for a pregnant woman is the gestational diabetes [4]. This type of diabetes is developed during pregnancy. The effect is leading to high blood pressure and too much protein in pregnant's urine. Several of the issues in case that the pregnant doesn't treat and control in time specific symptoms are: low blood glucose right after birth, breathing problems, a higher risk of dying before or soon after birth, may become overweight or obese or getting type 2 diabetes later on [4].

To overcome the previously mentioned issues it is vital to monitor very closely the pregnant woman and also keep her informed with what she must do in order to keep herself and baby as healthy as possible.

We integrated ObGyn Care, a mobile application for Android and IOS operating systems (OS) devices dedicated to the pregnant women by extending an existing application for the Obstetrics-Gynaecology Department (OGD) [5] accordingly with the Internet of things era. This new component of the OGD Information System is supporting better monitoring of patients, especially the ones with risk factors. ObGyn Care has the possibility to send the glycaemia, pulse, blood pressure, weight and contractions values to a database in the cloud. The physician has access to information from the office information system and monitors the pregnant women almost in real time. A component of the system allows the physician to see diagrams with the evolution of all parameters in different periods of time. The system sends notifications to the patient and to the physician if the value is not in the normal range.

## 1. Methods

The ObGyn Care application uses the related instruments and devices of the Internet of Things concept. It uses JQuery Mobile technology, HTML5, CSS3, Ajax and PHP making it available on smartphones and tablets, sending data to Windows Azure cloud, accessed by the physicians from their office application [6]. The other component is an extended module of the OGD Information System developed in Visual Studio.NET 2013 using ASP.NET pages and C# language. The associated devices are specific smart sensors and a smart phone.

The doctor can access the OGD Information System and can see what data was sent from the ObGyn Care app and if the system sent a notification the physician can contact the pregnant women and give some advices regarding the pregnancy, or if is critical to call the pregnant to come to the consultation.

The ECG smart sensor is a system composed by an Arduino Uno, ECG sensor, a Bluetooth port and has the possibility to send the data regarding blood pressure and pulse to a Smartphone which has the possibility to filter the data and sent to the cloud. This information may be later accessed by the OGD Information System and also by the ObGynCare App.

Figure 1 presents the system architecture.
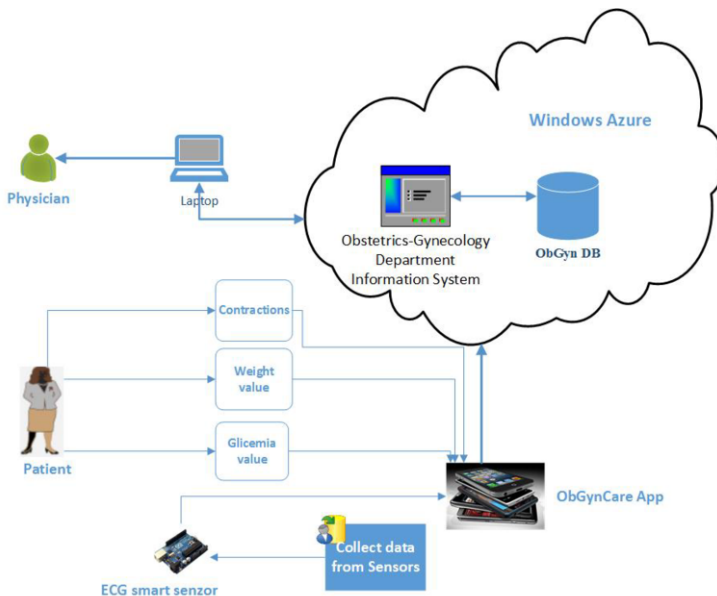
**Figure 1.** System architecture.

In order to ensure a continuity of care from the mobile application and the hospital department is needed a standardized communication, in this case is used HL7 CDA. HL7 CDA is a document mark-up standard that specifies the structure and semantics of clinical documents and it is represented as a XML document [7].

## 2. Results

ObGyn Care application is a mobile application which can be used on smartphones with IOS or Android OS. At the first use of the application, the patient must enter her name, her identification number, the pregnancy trimester and week. To other uses the application can be accessed automatically and also it computes and displays on the screen the pregnancy week and trimester.

The pregnant can enter her blood pressure, pulse values, weight and glycaemia in the ObGyn Care application, or to use the ECG smart sensor which measures and sends automatically data to the application. All of these values will be saved into a database on the cloud and can be accessed by the physician through the office application. The pregnant may set an alarm for a specific hour and this will alert her if the data are not entered in that day and recommends filling in. An alarm notifies the patient and the physician if the parameters values are not in the normal limits. Other functionality is that the patient can measure her contraction by using a button for start and stop when the contraction is happening, and can visualize the period between them. For a better communication with the physician, the application has a module where the patient can chat with the physician. The patient can visualize charts with the blood pressure, pulse, glycaemia and weight evolution on every week or for the entire trimester.

When they come to care in pregnancy, women are interested to search a lot of information related to recommended food, liquids as teas and alimentation that are best to consume during pregnancy [8]. A lot of things are restricted to be consumed during

pregnancy because some can do a lot of unwanted modification to the foetus. The proposed tool comes also to help pregnant women with general advice, being personalized for each one. For better information about pregnancy the application has an option where the patient can add meal plans. She can also share info on Facebook and change ideas with other pregnant women. Figure 2 presents six print screens from the ObGyn Care application, with the relevant functionalities.
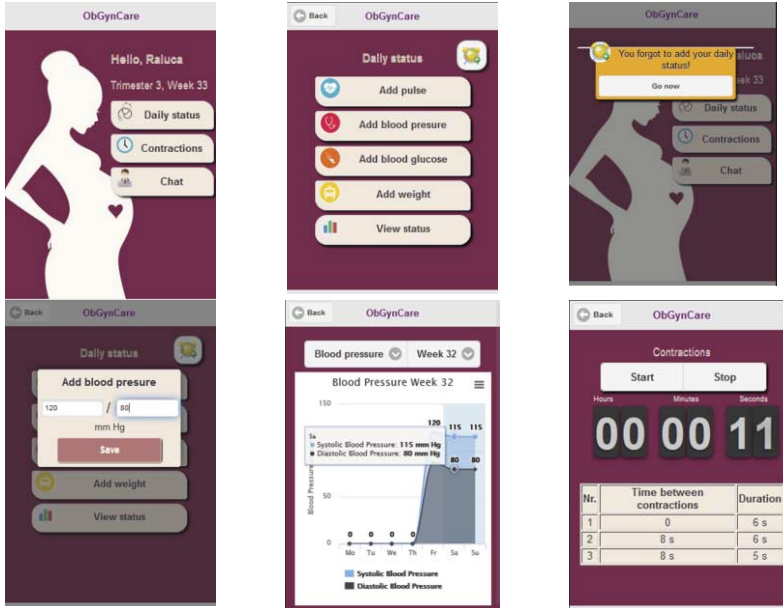


**Figure 2**. Screens from the ObGyn Care app.

The new component of the OGD Information System allows the physician to better monitor patients. Physicians have the possibility to visualize charts on specific week or trimester where he can follow the evolution of blood pressure, pulse, glycaemia or weight values, these parameters being selected in function of risk factors which can appear on the patient. Fig. 3 shows the blood pressure (A) and glycaemia (B) values evolution on 38th week for third trimester during pregnancy.
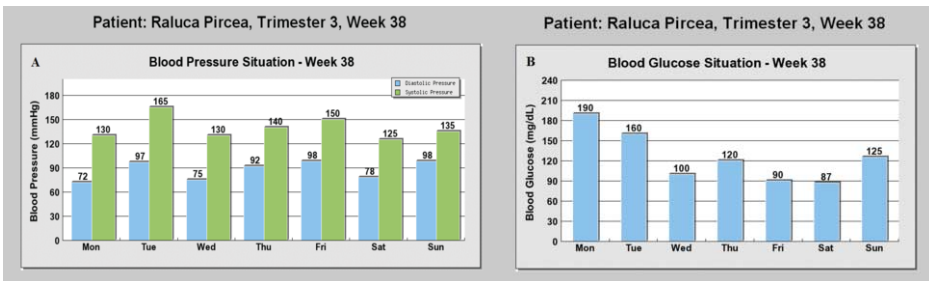


**Figure 3**. Screen from the Obstetrics-Gynaecology Department Information System
A. Monitoring the Blood Pressure on Week 38; B. Monitoring the Blood Glucose on Week 38.

## 3. Conclusion

Due to risk factors faced by pregnant patients, such as gestational diabetes, pre-eclampsia and obesity, a better monitoring is absolutely necessary. Collaboration from sides, patient and physician increases prevention of health problems of both the mother and the child, and reduces neonatal mortality. The ObGyn Care application works with two sorts of information: the first one is a general one about pregnancy and the second is a personalized part including the monitoring. The physicians, based on data from the mobile application stored in the cloud, view evolution diagrams of pulse, blood pressure, glycaemia and/or weight values for different periods of time, depending on the patient's risk factors.

The work described in this paper presents an integrated system for prevention and monitoring in Obstetrics-Gynaecology Domain composed by an application for IOS or Android OS mobile devices, called ObGyn Care, which supports pregnant women, and a new component of the OGD Information System, dedicated to the physicians for a better monitoring of patients, using the cloud, smart sensors, mobile devices and applications.

Testing started for the ObGyn Care app on 10 pregnant patients' mobile phones since 3 months. The feedback was registered using a quiz and free talks with patients. Testing in progress includes a usability test. The feedback from patients was positive, and the application is seen very useful for monitoring the health during pregnancy as well as for prevention purposes, ensuring in future mothers a feeling of trust and safety. The patients who suffer preeclampsia during pregnancy often have to go to the hospital to have their blood pressure measured. The feedback using the application is that the number of hospital visits can be reduced, the application ensuring a good monitoring of the health parameters and supporting easy communication with the physician. The application can monitor, among others, the blood pressure values in real time, sending them to the physician. If the values are not in normal range, the application will notify both the patient and the physician. In the future the app will be extended and uploaded on Google Play Store and App Store. As future possibilities due to the fact that the communication is a standardized one, the information may be sent to General Practitioner office and also to other departments (e.g. Neonatology department).

## References

[1] T. Chye, T. Tehng, T. Hseon, Practical Obstetrics and Gynaecology Handbook for the General Practitioner, World Scientific (2006).
[2] M. A. Leddy, M. L. Power, J. Schulkin, The Impact of Maternal Obesity on Maternal and Fetal Health, Rev Obstet Gynecol. 2008 Fall; 1(4): 170–178.
[3] A. Manaj, A. Rrugia, N. Manoku, The impact of preeclampsia in pregnancy, Prenat Med. 2011 Jan-Mar; 5(1): 19–22.
[4] Gestational diabetes, http://www.webmd.com/, Accessed in 20.10.2014
[5] M. Vida, L. Stoicu-Tivadar, O. Lupse, B. Blobel, E. Bernad, Models Supporting Development of Complex Information Systems in Healthcare. Case study: an Obstetrics-Gynaecology Department, *EJBI – Volume 9* (2013), Issue 1
[6] Fundamentals Introduction to Azure, https://azure.microsoft.com/en-us/documentation/articles/fundamentals-introduction-to-azure/, Accessed in 10.09.2015
[7] R.H.Dolin, L. Alschuler, C. Beebe, P.V.Biron, S.L. Boyer, D. Essin, E. Kimber et al, The HL7 Clinical Document Architecture, J Am Med Inform Assoc. 2001; 8:552 – 69
[8] A food guide for pregnant women, http://www.parents.com/, Accessed in 11.10.2014

# Adoption and Use of a Mobile Health Application in Older Adults for Cognitive Stimulation

Mobin YASINI[a,1] and Guillaume MARCHAND[a]

[a] *DMD santé, Research and development department, Paris, France*

**Abstract.** Serious games could be used to improve cognitive functions in the elderly. We evaluated the adoption of a new tablet application dedicated to cognitive stimulation in the elderly. The Stim'Art application offers various serious games to work different cognitive functions (memory, attention, concentration, etc.). The usage of fifteen older adults was followed for six months. The type of the game, the number of launches for each game, the time spent on each game, the difficulty level, the success rate and perceived well-being of users have been studied and compared at the end of the first and the sixth months. The participants have played half an hour per day on average. The average time of playing per day in the sixth month was significantly higher than the average time of playing during the first month *(p value $<7 * 10^{-4}$)*. The same result was found for the average number of game launches per day *(p value $<7 * 10^{-4}$)*. However, older people seem not to launch more difficult levels in the last month. The success rate at sixth months was significantly higher than the success rate at the end of the first month (*p value $<6.4 * 10^{-4}$*). Generally, seniors have had an improvement in their wellbeing score judged by themselves. Our study showed that the mobile application receives a good admission from users. The results are promising and can pave the way for improving cognitive function in the elderly patients. The use of tablets and the constitution of serious games in close cooperation with health professionals and elderly patients (the end user), are likely to provide satisfactory results to improve healthcare provided for elderly patients suffering from cognitive disorders.

**Keywords.** mHealth, Adoption, Cognitive Therapy, Mobile Application

## Introduction

The number of elderly dependents will quadruple by 2050. The economic impact of the long term care is likely to be very important [1]. No group of chronic diseases burdens the world more than mental illnesses [2]. Cognitive impairment is a major problem both for the elderly and for their family members and caregivers. The prevalence of cognitive impairment is increasing with the aging population and reaches up to 89% among older adults admitted in nursing homes [3–5]. Cognitive impairment has an adverse effect on psychosocial functioning of the elderly, affects the medical treatment and worsens the dementia. Therefore, it is important to pay more attention to the diagnosis and monitoring of cognitive status of residents in nursing homes. Apart from pharmacotherapy, that has a very limited effectiveness; various non-drug approaches

---

[1] Corresponding Author.

have been considered in the literature to improve neuropsychological and cognitive impairment symptoms. Music therapy, art therapy, [6,7] and cognitive stimulation due to digital activities and video games are some of these approaches [8–10].

Cognitive stimulation is an intervention for people with dementia which offers a range of enjoyable activities providing general stimulation for thinking, concentration and memory usually in a social setting [11]. This intervention could be done via the use of mobile health applications. Participation in the iPad intervention resulted in enhanced performance on cognitive constructs compared with control groups [12]. The advantage of digital approaches, including the use of tablet computers and mobile applications is the ability to use multiple approaches (art therapy, group activity, etc.). A digital approach allows a dematerialized and distant follow up of each user.

The main aim of this study was to evaluate the adoption of a new application dedicated to cognitive stimulation in the older adults.

## 1. Methods

### 1.1. The application

Stim'Art (standing for stimulation through electronic arts) is a mobile application (iOS and Android platform) that offers serious games on memory and brain training. A tracking site service is linked to the application to provide usage tracking per user. The application allows through various exercises to work different functions of the memory, in a playful way: working memory, attention, concentration, visual-spatial memory, etc. These exercises include general knowledge tests, puzzles of famous paintings, chronology of events, serious games on the music recognition and reaction speed exercises. Data entry burden, loss of interest and cost are the causes that make users stop using apps [13]. In the development of this app these factors have been considered.

### 1.2. Settings

We set up a pilot study to evaluate the adoption of Stim'Art application. This was done by setting 15 tablets for 15 seniors. A face-to-face workshop was run for each user to explain the application and the different games integrated therein. We asked our seniors to take control of their tablets and run the games included in the application whenever they want. We also explained our study and its purposes, asking them not to let others play with their tablet. The use of the application was tracked throughout the study via a computer system supported by the application editor. This system ensures the traceability for the tablet used (and therefore the user's sex and age), the launched serious game, the number of launches for each game, the time spent on the game, the difficulty level, and the success or failure to the played game (quitting the app without completing the game is considered as failure).

To assess the wellbeing of each user and its evolution, we calculated the average wellbeing perceived by each senior. Whenever the user wants to quit the application, it responds to a simple survey of well-being. The scores range from 1 (totally displeased) to 6 (delighted). The options to be selected by the user are totally displeased, dissatisfied, bored, fine, very well and delighted. The average score was calculated at the end of the first and the sixth month of the study. We used the Wilcoxon test for our paired data to compare the distribution of the information found in the two groups at

the end of the first and the sixth month. We also used McNemar test to check changes between two states (M1 and M6), in contingency tables. We finally used the Mann-Whitney test to compare the difference in improving the well-being between male and female users.

## 2. Results

There were 15 senior users (9 females and 6 males) and they were between 79 and 88 years old. Overall, our 15 users have played the application 78800 minutes for the period of 6 months. On average, each user has played half an hour (29.16 minutes) per day. Table 1 compares the average time spent during the first month and during the sixth month of use according to the difficulty levels.

**Table 1.** Average time spent per day per person in the first and sixth month of study according to the difficulty levels

| Average spent time (min/day/person) | Easy | Intermediate | Difficult | All levels |
|---|---|---|---|---|
| M1 | 12.37 | 7.99 | 4.7 | 25.06 |
| M6 | 12.54 | 12.4 | 13.47 | 38.41 |

The average time of use (game playing) per day for the sixth month was significantly higher than the average time during the first month ($p < 7 * 10^{-4}$). During the sixth month, the users played and allocated more time to more difficult levels than the easy levels in comparison with the first month. However, this difference was not statistically significant ($p < 0.999$). The various games of application were launched 4525 times by all users (1690 times during the first month and 2835 times during the sixth month). Table 2 compares the average number of game launches per day per person according to the difficulty levels.

**Table 2.** Average number of game launches per day per user in the first and sixth month of the study according to the difficulty levels

| Number of game launches/day/person | Easy | Intermediate | Difficult | All levels |
|---|---|---|---|---|
| M1 | 2.03 | 1.1 | 0.62 | 3.76 |
| M6 | 2.18 | 2.22 | 1.90 | 6.30 |

The average number of game launches per day during the sixth month was significantly higher than the average number of launches during the first month ($p < 7 * 10^{-4}$). This result confirms the previous result on the duration of use and shows that the more time passes, the more users adopt the app's games. We could not conclude that older adults launch more difficult levels in the last month ($p < 0.73$). Therefore, launching more difficult levels during the sixth month is due to the overall increase in launching the games. Success rates were calculated for the games launched in the first month and for games launched during the last month of the study. The results are displayed in table 3.

**Table 3.** The success rate (in percentage) at the end of the first and sixth months of the study according to the difficulty levels

| Success rate (%) | Easy | Intermediate | Difficult | All levels |
|---|---|---|---|---|
| M1 | 63.36 | 62.73 | 56.76 | 60.95 |
| M6 | 70.82 | 71.33 | 68.88 | 70.34 |

The success rate during the sixth months was significantly higher than the rate at the end of the first month ($p < 6.4 * 10^{-4}$). The result is distinctly the same for each level. In our observation, the difference in success rates between M1 and M6 increases based on the difficulty level. This means that older people may learn and would be able to successfully pass the challenges that improve their cognitive functions.

The average perceived well-being scores among the studied older adults are displayed in table 4.

**Table 4.** The average well-being scores for the first and last month of the study

| Senior ID | Sex | Wellbeing score M1 | Wellbeing score M6 |
|-----------|-----|--------------------|--------------------|
| 1 | M | 4.9 | 5.3 |
| 2 | M | 4.8 | 5.4 |
| 3 | F | 4.2 | 4.1 |
| 4 | F | 5.1 | 5.4 |
| 5 | M | 3.6 | 4.2 |
| 6 | F | 4.1 | 4.3 |
| 7 | M | 4.5 | 5.1 |
| 8 | F | 3.2 | 3.9 |
| 9 | M | 4.9 | 5.3 |
| 10 | M | 5.1 | 5.1 |
| 11 | F | 5.1 | 5.4 |
| 12 | F | 4.2 | 4.6 |
| 13 | F | 4.4 | 4.7 |
| 14 | F | 3.8 | 4.6 |
| 15 | F | 4.1 | 4.7 |

Apart from one user (number 3), all seniors had improvement in their perceived well-being score (*p value <0.001*). The average overall improvement was 0.41 (0.43 in men and 0.39 in women). However, there were no significant difference in improving the well-being between male and female users.

## 3. Discussion

In this study, we have set up a mobile health application that provides memory games adapted for seniors to improve their cognitive functions. When a senior runs the application, the selected game, difficulty level, time spent, success or failure, etc. are recorded by the system. We used these data during six months of using the application in order to assess the adoption of the application in the elderly as well as its effectiveness on the perceived well-being of users. The results were satisfactory and we found that users are never tired of playing with their tablet over time and that the average daily time spent on games increases with time. The number of launched games also drastically increases with time. These results show a good acceptability of the app's games that continues and improves with time. The success rate also improved over time. An improved success rate over time may imply a better game adoption by the elderly as well as an improved memory and cognitive abilities. The specific features of the application that explained the adoption could be the face-to-face workshop for the first use of the app, considering the ease of use and avoiding the adoption barriers in the development.

Our study is in line with other researches that have shown computer-assisted cognitive remediation is a therapeutic approach to enhance cognitive abilities [14,15]. The results of this study and the applications are promising and open new perspectives

for improving cognitive function in the elderly. Further studies with a larger number of users and considering various clinical factors co-related with cognitive stimulation, like clinical diagnosis and medication of each user are suggested to confirm these results. Using a neuropsychological scale to evaluate the effectiveness of these games to improve each cognitive function would be another perspective for future research. Studying the mood of the elderly and their participation in life would be of considerable interest. The participation of health professionals in the conception or development of health related apps is one of the criteria of creating reliable apps [16]. The use of tablets and the constitution of serious games in close cooperation with health professionals and the seniors (the end user), would lead to create tools that are more likely to succeed with good results once they are subject to a rigorous evaluation.

## References

[1] Mayston R, Guerra M, Huang Y, Sosa AL, Uwakwe R, Acosta I, et al. Exploring the economic and social effects of care dependence in later life: protocol for the 10/66 research group INDEP study. Springerplus. 2014;3.

[2] Fineberg NA, Haddad PM, Carpenter L, Gannon B, Sharpe R, Young AH, et al. The size, burden and cost of disorders of the brain in the UK. J Psychopharmacol. 2013;27(9):761-70.

[3] Hutsteiner P, Galler S, Mendoza MC, Klünemann HH. Prevalence of dementia in a rural nursing home population in Southern Germany. The European Journal of Psychiatry. 2013;27(3):174-84.

[4] Caracciolo B, Gatz M, Xu W, Pedersen NL, Fratiglioni L. Differential distribution of subjective and objective cognitive impairment in the population: a nation-wide twin-study. J Alzheimers Dis. 2012;29(2):393-403.

[5] Cahill S, Diaz-Ponce AM, Coen RF, Walsh C. The underdetection of cognitive impairment in nursing homes in the Dublin area. The need for on-going cognitive assessment. Age Ageing. 2010;39(1):128-31.

[6] Maratos AS, Gold C, Wang X, Crawford MJ. Music therapy for depression. Cochrane Database Syst Rev. 2008;(1):CD004517.

[7] Im ML, Lee JI. Effects of the art and music therapy on the depression and cognitive function of elderly. Technol Health Care. 2014;

[8] Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature. 2013;501(7465):97-101.

[9] Buschert VC, Friese U, Teipel SJ, Schneider P, Merensky W, Rujescu D, et al. Effects of a newly developed cognitive intervention in amnestic mild cognitive impairment and mild Alzheimer's disease: a pilot study. J Alzheimers Dis. 2011;25(4):679-94.

[10] Oei AC, Patterson MD. Enhancing cognition with video games: a multiple game training study. PLoS ONE. 2013;8(3).

[11] Woods B, Aguirre E, Spector AE, Orrell M. Cognitive stimulation to improve cognitive functioning in people with dementia. Cochrane Database Syst Rev. 2012;2:CD005562.

[12] Chan MY, Haber S, Drew LM, Park DC. Training Older Adults to Use Tablet Computers: Does It Enhance Cognitive Function? Gerontologist. 13 juin 2014;

[13] Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth. 2015;3(4).

[14] Siegle GJ, Ghinassi F, Thase ME. Neurobehavioral Therapies in the 21st Century: Summary of an Emerging Field and an Extended Example of Cognitive Control Training for Depression. Cogn Ther Res. 2007;31(2):235-62.

[15] Naismith SL, Redoblado-Hodge MA, Lewis SJG, Scott EM, Hickie IB. Cognitive training in affective disorders improves memory: A preliminary study using the NEAR approach. Journal of Affective Disorders. 2010;121(3):258-62.

[16] Yasini M, Marchand G. Mobile Health Applications, in the Absence of an Authentic Regulation, Does the Usability Score Correlate with a Better Medical Reliability? Stud Health Technol Inform. 2015;216:127-31.

# Positioning Commercial Pedometers to Measure Activity of Older Adults with Slow Gait: At the Wrist or at the Waist?

Frederic EHRLER[a,1], Chloe WEBER [b], and Christian LOVIS [a,c]

[a] *Division of medical information science, University Hospitals of Geneva, Geneva, Switzerland*
[b] *EPFL, Lausanne, Switzerland*
[c] *University of Geneva, Geneva, Switzerland*

**Abstract.** Sedentary behaviour is a major risk factor for chronic disease morbidity and mortality in aging. Measuring people activity through devices such as pedometers is a recognized intervention to motivate them for more physical activity. However, the feedback provided by these devices must be accurate in order to avoid overtraining and keep users' motivation alive. If the accuracy of pedometers has been validated for healthy people, their lack of accuracy for elderly people walking at slower pace has been reported in several studies. The emergence on the consumer's market of new devices that can be worn indifferently at the wrist or at the waist raises once more this concern. In order to evaluate whether pedometers' location influences their accuracy, we have tested three pedometers at different locations, and for several paces in a comparative study. Beyond confirming the decrease of pedometers' accuracy with speed reduction, our study reveals that pedometers should be worn at the waist rather than at the wrist. This leads us to recommend wearing pedometers at the waist when monitoring population with reduced mobility.

**Keywords.** Independent Living, Frail Elderly, mHealth, Walking

## Introduction

There is no more need to demonstrate the importance of physical activity for older people for the prevention of disease, maintenance of independence and improvement of quality of life [1]. If physical activity is particularly important among older adults, statistics unfortunately reveal a constant decrease of activity after 65 years. A cheap and efficient intervention to motivate people activity is the use of pedometers. Indeed, the goal setting theory taught us that measuring one's activity, setting suitable goals, and receive positive feedback are motivating factors for increasing physical activities [2]. However, the efficacy of these devices to motivate people may be influenced by the quality of the feedback provided after exercising. In case this feedback is inaccurate, people may disengage from the goal pursued.

If most pedometers have demonstrated their ability to accurately count footsteps of healthy adults walking at their preferred speed, many studies have raised concerns

---

[1] Corresponding Author.

about the efficacy of these devices to measure activity of older adults walking at a slower pace. Frail elderly such as diabetic, obese, or patients suffering from heart failure often walk at a slower pace (around 0.6 m/s [3], and as low as 0.4 m/s for community ambulation [4]). At such pace, many pedometers show a decrease of accuracy with relative errors going from 30% to 60% [5]–[7].

While, until recently, most pedometers were worn at the waist, the new generation of pedometers entering the market is more versatile and can also be worn at the wrist as a bracelet. The possibility to wear pedometers at different locations sets the question of the location that ensures the best accuracy. In order to respond to this question, we have conducted a comparative study to explore the impact of the location of several pedometers on their accuracy at different walking speeds.

## 1. Methods

A comparative study has been set up to compare three different pedometers worn at two different locations (waist and wrist) and at four walking speeds.

### 1.1. Instruments

Three commercial pedometers, iHealth activity monitor, Withings pulse O2, and Misfit Shine, released in 2013, have been selected for our experiment. These three devices have been selected because they can be worn either at the waist (attached to the belt) or at the wrist (as a bracelet), but also for their suitability to be used in a medical scenario. Indeed, iHealth and Withings are brands that offer a variety of medical wearables that allow monitoring a large spectrum of health parameters. Misfit has been added for its very small size, and its potential acceptance by elderly that are particularly sensitive to stigmatization.

### 1.2. Population

Based on a previous study demonstrating that kinematic and kinetic graph patterns did not show specific differences between elderly and young subjects, we decided to enrol healthy adults [8].

### 1.3. Experimental setting

The walking speed range selected for our experiment is [0.4 m/s - 0.8 m/s]. The lower limit has been chosen based on a classification of walking handicap in the stroke population proposed by Martin and al. [5] that recognizes 0.4 m/s as the minimum speed for elderly able to walk outside of their house. If 1.4 m/s is recognized as the preferred walking speed of humans, we have set our upper limit at 0.8 m/s since it is the limit determined for normal speed.

In order to control the walking speed of participants, a technique combining the regulation of the footstep frequency by a metronome and the limitation of the footstep length by a rope has been employed. Based on a study exploring the relationship between length and frequency of the footsteps of physically active women, we have chosen a ratio of 0.55 between the footstep length and the cadence [9]. Consequently,

for each targeted speed, the footstep length can be determined using this ratio in the equation (1) expressing the relation between step length, speed, and cadence.

$$stepsLength = speed/cadence \qquad (1)$$

$$stepsLength(cm) = \sqrt{speed * stepsLength/cadence} = \sqrt{speed(m/s) * 3300} \qquad (2)$$

Furthermore, the relationship between the speed, the cadence, and the step length can be also derived from the equation (1) up to the following equation:

$$Cadence\left(\frac{steps}{min}\right) = \frac{speed\left(\frac{m}{s}\right) * 6000}{stepsLength\,(cm)} \qquad (3)$$

Based on the equations (2) and (3), the different settings of the experiment are calculated and presented in Table 1.

**Table 1.** Relation between speed, step length, and cadence

| Speed | 0.4 m/s | 0.6 m/s | 0.8 m/s |
|---|---|---|---|
| Step length | 36 cm | 44 cm | 51 cm |
| Cadence | 66 steps/min | 82 steps/min | 93 steps/min |

### 1.4. Course of the study

The experiment took place on a running track where distances were delimited using marks on the ground. For each of the pedometers selected in our study, participants wore one device attached to the belt and another as a bracelet. In a first stage, participants were requested to walk 200 m at their preferred speed. In a second phase, participants were invited to walk 100 m at a speed constrained by a rope linking their legs and a pace given by a metronome.

All walks were videotaped. The videos have been watched afterward to count the actual number of footsteps done by each participant. In order to ensure the quality of the measures, the count has been double-checked by two different persons for each walk. When the two counts were different, the process was restarted until we get two identical figures. This measure was taken as the reference value acting as baseline to be compared with the count provided by pedometers. The mean relative error was calculated by averaging, over all participants, the absolute differences between the count done by observers (actual count) and the count given by the pedometer.

## 2. Results

Twenty-one persons (12 women and 9 men) participated to the study. Participants' age ranged from 24 to 80 year-old with a median at 30 year-old.

The mean relative counting errors averaged over all participants at each speed and for each brand are presented in **Table 2**. Additionally, the results are presented graphically as a tornado graph (Figure 1) in order to give a more visual insight. The tornado graph provides on a single graph a vision of the influence of the speed and the position on the error rate. The left side represents the errors of pedometers worn at the wrist and the right side represents the errors of pedometers worn at the waist. Each line

represents a single pedometer and each histogram branch presents the cumulated errors in percentage at each speed.

**Table 2.** Absolute mean relative errors in percentages

| Brand | Location | Preferred speed | 0.8 m/s | 0.6 m/s | 0.4 m/s |
|-------|----------|-----------------|---------|---------|---------|
| iHealth Activity | Wrist | 10.21 | 14.79 | 26.97 | 62.64 |
| | Waist | 0.55 | 5.12 | 16.29 | 56.45 |
| Withings Pulse O2 | Wrist | 14.37 | 30.20 | 64.07 | 88.51 |
| | Waist | 0.87 | 18.27 | 80.92 | 99.34 |
| Misfit Shine | Wrist | 37.16 | 55.08 | 40.18 | 55.90 |
| | Waist | 39.05 | 40.93 | 49.87 | 70.51 |

Regarding the brand, the results revealed that, except at 0.4 m/s, the iHealth pedometer has better accuracy than those of the other brands independently of the position. The worst model is Misfit for which the accuracy never rises above 65%.

The choice of the pedometers' location is less obvious. Indeed, if wearing an iHealth pedometer at the waist produces fewer errors than wearing it at the wrist independently of the speed, for Withings and Misfit, the wrist becomes the most appropriate position when the speed drops under 0.6 m/s.

Finally, a comparison of the accuracy regarding to the speed reveals a decrease of accuracy going along with the decrease of speed. Whereas, at natural speed, the error was inferior to 1% for iHealth and Withings pedometers when worn at the waist, the error rises quickly with speed decrease until reaching almost 100% for Withings and more than 50% for iHealth at 0.4 [m/s].
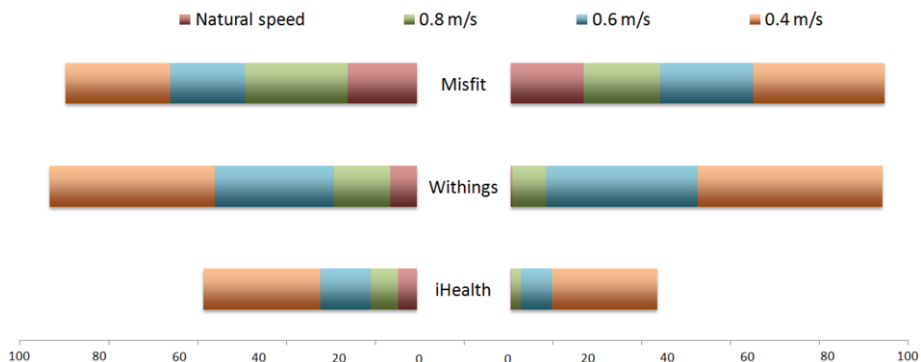


**Figure 1.** Tornado graph of the relative errors of pedometers for the wrist (left) and waist (right) locations

## 3. Discussion

The results evidence a decline of pedometers' accuracy associated to the reduction of the walking speed. This confirms the observations done in several previous studies: Schmidt-Trucksäss et al. reported errors close to 24% at 0.66 [m/s] and around 9% at 0.83 [m/s] when testing pedometers on chronic heart failure patients [3]. Martin et al., reported errors going from 9% for all devices at self-selected speed to 56% at 50 steps/min [5].

This loss of accuracy at a slow pace can be due either to the quality of the sensors inside the pedometers or to the detection algorithm. Indeed, most algorithms identify

footsteps based on the vertical acceleration. Since the vertical acceleration diminishes accordingly to the speed of the walk, it is harder to detect every footstep at a slow pace. Unfortunately, as most brands keep their algorithm's specification confidential, it is not possible to draw conclusion based on these factors.

Finally, if providers advertise these new devices as versatile, with the possibility to wear them at various locations, we have observed that the accuracy clearly depends on the location they are worn. The lower accuracy of pedometers worn at the wrist is certainly due to the higher complexity of the acceleration pattern produced at this location. Indeed, at the wrist, the acceleration recorded is a combination of the vertical acceleration generated by the walk and the acceleration produced by arm movements.

## 4. Conclusion

Monitoring activity with pedometers has the potential to motivate people to increase physical activity. This can be particularly beneficial for frail elderly. Unfortunately, the poor performance of consumer market pedometers at low speed makes us doubt about the real effectiveness of these devices for such population.

As no validation authority controls the accuracy of consumer market pedometers, it is pretty complicated to trust these devices without performing real-world tests. Moreover, most websites that propose pedometers' comparisons rather base their reviews on aesthetic features than on the actual accuracy of devices.

If the lack of accuracy at low speed has been already demonstrated in several studies, our work emphasizes that the loss of accuracy can be further worsened if pedometers are worn at the wrist, which reduces the interest of this new generation of devices.

## References

[1] F. Sun, I. J. Norman, and A. E. While, "Physical activity in older people: a systematic review.," *BMC Public Health*, vol. 13, no. 1, p. 449, 2013.

[2] C. E. Tudor-Locke, "Taking steps toward increased physical activity: Using pedometers to measure and motivate," *Pres. Counc. Phys. Fit. Sport. Res. Dig.*, vol. 3, no. 17, pp. 3–10, 2002.

[3] M. Jehn, a Schmidt-Trucksaess, T. Schuster, H. Hanssen, M. Halle, and F. Koehler, "Pedometer Accuracy in Patients with Chronic Heart Failure," *Int. J. Sports Med.*, vol. 31, no. 3, pp. 186–191, 2010.

[4] M.-V. Castell, M. Sánchez, R. Julián, R. Queipo, S. Martín, and A. Otero, "Frailty prevalence and slow walking speed in persons age 65 and older: implications for primary care.," *BMC Fam. Pract.*, vol. 14, no. 1, p. 86, 2013.

[5] J. B. Martin, K. M. Krč, E. a Mitchell, J. J. Eng, and J. W. Noble, "Pedometer accuracy in slow walking older adults.," *Int. J. Ther. Rehabil.*, vol. 19, no. 7, pp. 387–393, Jul. 2012.

[6] J. Lauritzen, A. Muñoz, J. Luis Sevillano, and A. Civit, "The usefulness of activity trackers in elderly with reduced mobility: A case study," *Stud. Health Technol. Inform.*, vol. 192, pp. 759–762, 2013.

[7] M. Marschollek, M. Goevercin, K.-H. Wolf, B. Song, M. Gietzelt, R. Haux, and E. Steinhagen-Thiessen, "A performance comparison of accelerometry-based step detection algorithms on a large, non-laboratory sample of healthy and mobility-impaired persons.," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2008, pp. 1319–1322, 2008.

[8] W. S. Kim and E. Y. Kim, "Comparing self-selected speed walking of the elderly with self-selected slow, moderate, and fast speed walking of young adults," *Ann. Rehabil. Med.*, vol. 38, no. 1, pp. 101–108, 2014.

[9] A. Zijlstra, E. D. de Bruin, N. Bruins, and W. Zijlstra, "The step length-frequency relationship in physically active community-dwelling older women.," *Eur. J. Appl. Physiol.*, vol. 104, no. 3, pp. 427–34, Oct. 2008.

# Towards a Wireless Smart Polysomnograph Using Symbolic Fusion

Adrien UGON[a, 1], Brigitte SEROUSSI[b,c], Carole PHILIPPE[d], Jean-Gabriel GANASCIA[a], Patrick GARDA[a], Karima SEDKI[b], Jacques BOUAUD[b,e] , and Andrea PINNA[b]

[a] *Sorbonne Universités, UPMC Univ Paris 06, UMR 7606, LIP6, Paris, France*
[b] *Sorbonne Universités, UPMC Univ Paris 06, INSERM, Sorbonne Paris Cité, Université Paris 13, LIMICS, UMR_S 1142, Paris, France*
[c] *AP-HP, Hôpital Tenon, Département de Santé Publique, Paris, France*
[d] *AP-HP, Hôpital Pitié-Salpêtrière, Unité Pathologies du sommeil, Paris, France*
[e] *AP-HP, DRCD, Paris, France*

**Abstract.** Polysomnography is the gold standard test for sleep disorders among which the Sleep Apnea Syndrome (SAS) is considered a public health issue because of the increase of the cardio- and cerebro-vascular risk it is associated with. However, the reliability of this test is questioned since sleep scoring is a time-consuming task performed by medical experts with a high inter- and intra-scorers variability, and because data are collected from 15 sensors distributed over a patient's body surface area, using a wired connection which may be a source of artefacts for the patient's sleep. We have used symbolic fusion to support the automated diagnosis of SAS on the basis of the international guidelines of the AASM for the scoring of sleep events. On a sample of 70 patients, and for the Apnea-Hypopnea Index, symbolic fusion performed at the level of sleep experts (97.1% of agreement). The next step is to confirm these preliminary results and move forward to a smart wireless polysomnograph.

**Keywords.** Sleep Apnea Syndromes, Polysomnography, Wireless Technology, Artificial Intelligence, Symbolic Fusion

## Introduction

Sleep Apnea Syndrome (SAS) is a disorder characterized by repeated respiratory disturbances occurring during sleep. Estimations of SAS prevalence range from 3 to 7% [1]. The early diagnosis of SAS is a public health problem since numerous studies have identified that the disease is associated with increased cardio- and cerebro-vascular morbidities [2,3]. Besides, markedly impaired quality of life [4], increased risk of road traffic [5], and work accidents [6] are also acknowledged consequences of SAS.

Polysomnography is the gold standard diagnostic test for sleep disorders among which SAS. It consists in recording different physiological signals during a whole night. Resulting curves (Figure 1.a) are analysed by sleep experts who assess sleep cycles, highlight abnormal events, and relate them to each other by cause and effect relationships. Some curves are used to stage sleep, others to detect apneas, the last ones are used to

---

[1] Corresponding Author.

recognize physiological events that occur as consequences of apneas (desaturations, legs movements). The device used to record physiological signals is a polysomnograph. It comprises a central box to which commonly more than 15 sensors are connected by wired connections (Figure 1.b). Data are then stored and processed on a standard computer.
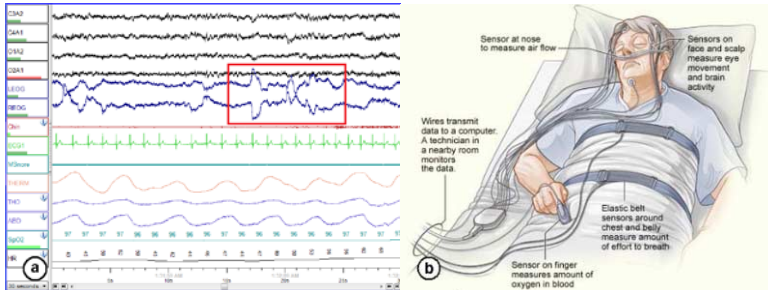


**Figure 1.** A polysomnograph generates curves (a) from sensor-collected data (b).

The numerous cables used are a source of artifacts for the patient's sleep, which may have an impact on the reliability of the diagnosis. Moreover, sleep scoring is currently performed by sleep experts. This is a tedious and time-consuming task. The inter and intra scorers variability is a real issue, although partially solved by the publication of international guidelines [7]. However, it is still acknowledged that scoring practices need to be improved [8].

We propose two ways to improve the reliability of polysomnography. First, using wireless sensors should help to provide the patient a sleep environment similar to his/her usual sleeping conditions. Second, automatically processing data from sensors with appropriate methods should guarantee the reproducibility of the scoring. Most of marketed digital sleep devices include automatic scoring functionalities, but they are not fully satisfying and sleep physicians do not use them. Research work is being conducted to automate the scoring [9]. Most of this work uses signal-processing methods to extract patterns that are then injected into automatic classification algorithms to recognize abnormal events or sleep stages. However, these methods are not efficient in diagnosing SAS because they are not able to take into account the semantic representations experts use to identify the relationships between the abnormal event and all related events. Unlike purely numerical methods, symbolic fusion is a method used in artificial intelligence that has proven its efficiency to process low level data (raw data coming from heterogeneous sources) and reach a high level, through the management of a semantically-tagged information. This relies on formally represented knowledge and reasoning processes through different abstraction levels [10].

In this paper, we describe the first steps of a project that aims at developing a wireless smart polysomnograph through the use of remote sensors able to record physiological parameters, and the application of symbolic fusion to support the diagnosis of SAS.

## 1. Methods

We used the set of widely accepted recommendations defined by the American Academy of Sleep Medicine (AASM) for the scoring of sleep and of associated events on

polysomnographic curves [11]. Most of the events are defined by the presence of concomitant or lightly delayed observations on the curves. Scored events are used to generate several indexes used for diagnosis. Among them, the Apnea-Hypopnea Index (AHI), defined as the average number of apneas and hypopneas per hour of sleep, is used to assess the existence and severity of a SAS.

Information fusion is defined by Henrik Boström et al. as "the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human or automated decision making" [12]. Symbolic fusion uses a formal representation of concepts and fusion strategies. Symbolic concepts are intelligible for experts, which allows to understand how decisions are taken, and helps improve the system by adding or correcting rules.

Figure 2 displays an example of the application of symbolic fusion to stage rapid eye movement (REM) sleep stage. Raw input data come from six different sensors, three electroencephalograms (EEG) for brain activity, two electrooculograms (EOG) for eyes movements and one electromyogram (EMG) for muscle tone. Three layers – data, feature, and decision – are crossed to process these signals and decide whether the treated portion of signal meets the criteria of REM sleep stage. In the "feature" layer, parameters are introduced to reflect the criteria defined in the guidelines.
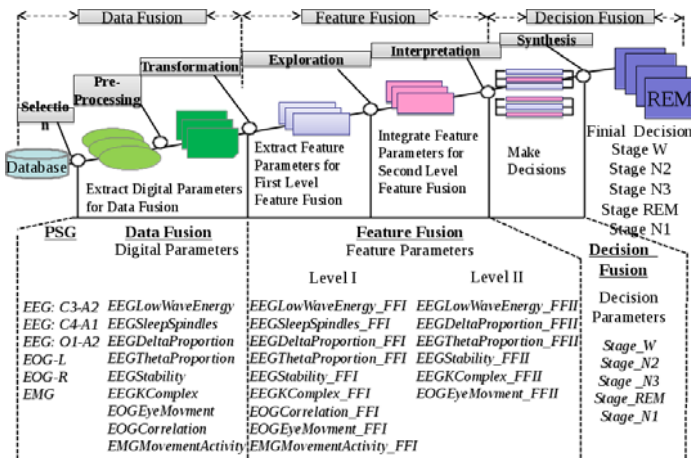


**Figure 2.** Flowchart of data through the different layers of symbolic fusion to stage REM sleep stage.

Symbolic fusion is implemented by a distributed processing architecture where several abstraction layers are built one above the other, from the sensor to the central unit, to satisfy a low power consumption of autonomous wireless sensors and the correct execution of the expected task (polysomnography recording). Low layers treatments can be processed directly on sensors; upper layers treatments are then processed on a remote computer. We first decomposed sleep into stages, then extract respiratory events for each of these stages. Both steps involve different sub-tasks to recognize and extract relevant patterns and events.

## 2. Results

We present preliminary results of the use of symbolic fusion to automate SAS diagnosis. Data were collected from 70 patients of the Tenon Hospital (AP-HP), in Paris, France. The sample was composed of 24 men and 46 women, aged 55 ± 12.5 years, with a BMI of 29.9 ± 6.0 kg/m². Data were recorded using the Embla system®; scoring was performed with the Somnologica Software™ (Resmed).

Results produced by symbolic fusion have been compared with those provided by an expert scorer. Figure 3 displays the results for the AHI with data given by the expert, resp. the symbolic fusion, presented on the x-axis, resp. the y-axis. The different ranges of disease severity are represented by colored rectangles. The result is "perfect" when the expert and the symbolic fusion agree, i.e. the point is located on the median axis. It is satisfying when it is located within the same rectangle. It is acceptable when the AHI is overvalued (located above the rectangle), but unsatisfactory when the point is located below the rectangle. Indeed, when the AHI is above a given threshold, a treatment is required. The adaptation of the treatment depends on other parameters and should be decided by a physician. The analysis of the data provided two cases in this later situation. The two points are surrounded by a red circle. The performance of the symbolic fusion was measured by an agreement ratio of 97.1% (68/70).
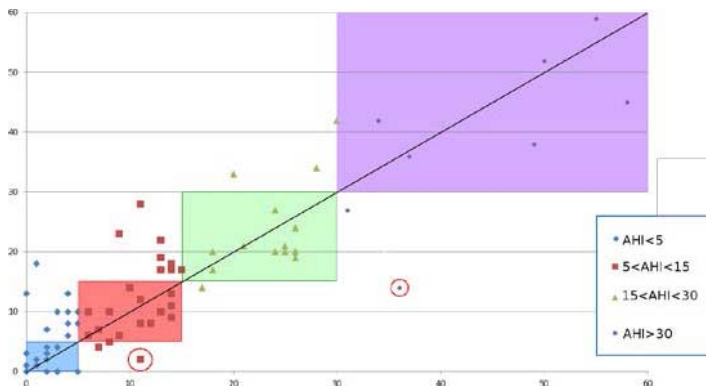


**Figure 3.** Comparison of results provided by the expert and the symbolic fusion process.

## 3. Discussion

Symbolic fusion gave satisfying results to support the automatic scoring of polysomnographic data except in two cases in a sample of 70 patients. In the first case (red square), the event lasted exactly 10 seconds, which was the threshold for an event to be considered. In the other case, the acquired signal was very noisy, which seems to have disrupted the automatic analysis.

Very few works completely automatize the support for SAS diagnosis. Most of them focus on only one step, for example on staging sleep or on recognizing sleep events like apneas or arousals. Diego Álvarez Estévez has proposed in his PhD thesis a method based on fuzzy inference reasoning [13]. Although fuzzy inference reasoning and symbolic fusion provide similar advantages, fuzzy inference reasoning can be used in several consecutive abstraction levels. It is also capable of representing and processing

imprecision. Results can be expressed in a human-like understandable way through the use of linguistic terms, which facilitates the explanation of results. However, the reasoning process remains numeric to estimate the degree of membership of a point to a fuzzy set, while symbolic fusion focuses on the use of semantically formalized concepts, and on the application of guideline-based explicit rules.

In future work, we'll consider adding new rules to evaluate new indices for the analysis of sleep disorders in order to improve symbolic fusion in case of noisy sources, and also to consider the clinical context of patients (to support the interpretation of the indexes computed by the events scored on polysomnographic curves, e.g. the AHI). The other dimension of improvement would be to go for a smart wireless polysomnograph implemented using Field Programmable Gate Array (FPGA) circuit.

## Acknowledgement

## References

[1] Marin JM, Carrizo SJ, Vicente E, Agusti AGN. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* **365**(9464), 104653, 19 March 2005.

[2] Punjabi NM. The epidemiology of adult obstructive sleep apnea. *Proc Am Thorac Soc*, 2008 Feb. 15;**5**(2), 13643.

[3] Arzt M, Young T, Finn L, Skatrud JB, Bradley TD. Association of sleep-disordered breathing and the occurrence of stroke. *Am J Respir Crit Care Med*, **172**(11), 1447–1451, Dec 2005.

[4] D'Ambrosio C, Bowman T, Mohsenin V. Quality of life in patients with obstructive sleep apnea: Effect of nasal continuous positive airway pressure—a prospective study. *Chest*, **115**(1), 123–129, 1999

[5] Terán-Santos J, Jimenez-Gomez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. *New England Journal of Medicine*, **340**(11), 847–851, 1999

[6] Uehli K, Mehta AJ, Miedinger D, Hug K, Schindler C, Holsboer-Trachsler E, Leuppi JD, Künzli N. Sleep problems and work injuries : A systematic review and meta-analysis. *Sleep Medicine Reviews*, **18**(1), 61–73, feb 2014

[7] Kuna ST; Benca R; Kushida CA; Walsh J; Younes M; Staley B; Hanlon A; Pack AI; Pien GW; Malhotra A. Agreement in computer-assisted manual scoring of polysomnograms across sleep centers. *Sleep* 2013;**36**(4):583-589.

[8] Penzel T; Zhang X; Fietze I. Inter-scorer reliability between sleep centers can teach us what to improve in the scoring rules. *J Clin Sleep Med* 2013;**9**(1):89-91.

[9] Alvarez-Estevez D, Moret-Bonillo V. Computer-Assisted Diagnosis of the Sleep Apnea-Hypopnea Syndrome: A Review. *Sleep Disorders*, **2015**(8), 1-33.

[10] Dasarathy BV. Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proc IEEE* 1997, **85**(1), 2438.

[11] Berry RB, Brooks R, Gamaldo CE, Harding SM, Lloyd RM, Marcus CL and Vaughn BV for the American Academy of Sleep Medicine. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.2. www.aasmnet.org. Darien, Illinois: American Academy of Sleep Medicine, 2015.

[12] Boström H, Andler SF, Brohede M, Johansson R, Karlsson A, Van Laere J, et al. On the Definition of Information Fusion as a Field of Research. Informatics Research Centre, University of Skövde,. Tech. Rep. HS-IKI-TR-07-006, 2007.

[13] Estévez DÁ. Diagnosis of the sleep apnea-hypopnea syndrome: a comprehensive approach through an intelligent system to support medical decision. PhD thesis, Univ de La Rioja, Spain, October 2012;

This page intentionally left blank

# Societal Dimensions of the IoT

This page intentionally left blank

# Risk and the Internet of Things: Damocles, Pythia, or Pandora?

Chris SHOWELL[a,1]

[a] *eHealth Services Research Group, University of Tasmania, Australia*

**Abstract.** The Internet of Things holds great promise for healthcare, but also embodies a number of risks. This analysis suggests that the risks are as yet poorly delineated (having features in common with the oracle Pythia, and with Pandora and her box), and that adopting the precautionary principle is appropriate.

**Keywords.** Risk; Privacy; Computer Security; Internet

## Introduction

The Internet of Things (IoT), is " an infrastructure of interconnected objects, people, systems and information resources together with intelligent services to allow them to process information of the physical and virtual world and react."[1] Smith and Erickson envision an IoT connected future in which "...the myriad embedded devices magically enhance our living environments, adjusting lights, temperature, music, medication, fuel flow, traffic lights, and elevators."[2]

For healthcare, the IoT represents an emerging sociotechnical environment which will change the way in which information and communication technology (ICT) is used, and is likely to provide a range of functions, including diagnosis, monitoring, treatment, and ambient assisted living. However, with any tantalising promise comes a range of poorly delineated or unknown risks. The IoT is a rapidly evolving ecosystem rather than a purpose designed healthcare environment. As Smith and Erickson note "…the IoT will probably grow organically, a global mashup of heterogeneous components with no top-down set of principles determining its emergent behaviour."[2] Each patient is likely to have an unpredictable, and uncurated collection IoT devices, including direct-to-consumer items, 'professional' devices recommended or provided by healthcare providers, and mundane domestic items. Because the management of this environment will be complex, it is appropriate to consider the threats and risks associated with these developments as examples of the *systemic* risks normally associated with threats to the environment. A framework for the analysis of systemic risk developed by the German Advisory Council on Global Change (WBGU) considers risk and damage, as well as the impact of uncertainty, ubiquity, persistency, reversibility, delay effects, violation of equity, and the potential for individual and community mobilisation.[3] This paper seeks to provide a broad, discursive overview of some of the challenges, which the IoT presents from the perspective of systemic risk.

---

[1] Corresponding Author.

## 1. Methods

The pace of development in the IoT is rapid and accelerating. Technical and financial barriers to the entry of new devices are low, and health-focused IoT devices and systems reside in an environment already crowded with such mundane items as refrigerators, toasters, and light globes. Taken together, these factors mean that risk likelihood and damage extent are not easy to estimate

In consequence a loosely structured approach has been taken to review security and privacy considerations, threats and risks, find examples of IoT incidents, and consider types of corporate behaviour which could have an effect on the use of the IoT for healthcare. A search in PubMed using the term "Internet of Things" returned 170 items. On review, all but 13 were discarded as not relevant, and only two described IoT risks. A comparable search in Google Scholar provided another three publications. The small number of retrieved publications may be a reflection of the rapid advance of the field. Searches in Google and Yahoo! for "Internet of Things" and "risk" or "fraud" yielded another 62 items (including blogs, and magazine and newspaper articles). These were reviewed to identify reported remote attacks, including on IoT devices, as well as likely avenues for future interference with IoT-based health-related processes.

## 2. Results

### 2.1. Security

Several challenges to health IoT were identified. Kozlov et al [4] detail a number of risks, including the user losing control of the device, interception of sensor data during transfer to a health provider, the unavoidable trade-off between usability and security, a potential misuse of personal identifiers following eavesdropping, theft of data from a data aggregation system, and targeted denial of service attacks. Identity theft, by impersonating the device's IP address (IP spoofing) or MAC address (ARP spoofing) was also reported as a risk.[5] Reback and Costello [6] warn that "IoT development could increase security vulnerabilities at both the individual and systemic levels. Part of the problem is that while a variety of new "things" are being connected to the Internet, the manufacturers of these objects may not have either the experience or expertise to implement appropriate security safeguards."

### 2.2. Privacy

Ensuring the privacy of health data is a matter of concern for a significant proportion of patients. Papoutsi et al [7] found that 79% patients had concerns about the security of their data within an electronic health record. Typically, the privacy protection offered to end-users of commodity technology is provided on the basis of 'notice and consent' – users are asked to read extensive details about how their data will be used, and then consent to the terms that are offered. Since many IoT devices lack a user interface, it is not clear how the privacy conditions will be agreed to.

Baby monitors also present a threat to privacy. Weaknesses in the camera software have allowed hackers to remotely access a number of Internet-connected baby monitors, and communicate verbally with the monitored infants.[8]

"By connecting to Wi-Fi, these so-called Internet of Things (IoT) devices allow access from wherever the owner is in the world, but…these devices are often so poorly secured, it takes little effort for a hacker to gain access." [9]

## 2.3. Maintenance and update

Most software driven devices require periodic updates to ensure that they continue to operate appropriately, and are protected against vulnerabilities. IT professionals may find it challenging to determine whether a computer is affected by a newly identified vulnerability; updating IoT device operating systems and applications will be onerous for consumers with sound technical skills, and near impossible for those without. "Left unperturbed, the commercial/industrial IoT sector might prove to be the Windows XP of the IoT, full of homogeneous badness that won't go away and persisting through an unwillingness to embrace improved, potentially safer systems."[2]

## 2.4. Hacking examples

There are already several examples of hacking attacks affecting IoT devices, as well as larger Internet-connected healthcare appliances. Two security researchers remotely accessed nearly 70,000 medical systems within a large US healthcare organisation, including "…21 anaesthesia, 488 cardiology, 67 nuclear medical, and 133 infusion systems, [and] 31 pacemakers".[10] In 2015 the US Food and Drug Administration issued a warning that a model of infusion pump in common use in hospitals was vulnerable to cyber attack, and recommended withdrawing the pumps from use.[11]

When Dick Cheney's defibrillator was replaced in 2007, the wireless interface was inactivated to minimise the risk of assassination.[12] In July 2015 Charlie Miller and Chris Valasek demonstrated their ability to control the radio, climate control and accelerator of a Jeep Cherokee by remotely accessing the car's control systems.[13]

## 2.5. Corporate behaviour

There have been several reports of actions and strategies of dubious ethical standing which are seen as 'business as usual' for large corporations.

*Corporations may knowingly lie about or obscure the function of a device.* Volkswagen manipulated the software in its Jetta diesel engines to disable the emission control system except when a vehicle appeared to be undergoing laboratory emissions testing. [14] Samsung issued a statement denying that the power saving "motion lighting" feature in their televisions had been specifically designed to give misleading results in IEC power consumption tests, although that was an effect of the feature.[15] The everyday operation of late model John Deere tractors is managed by proprietary software that the manufacturer has protected by copyright, effectively preventing maintenance by any but company-licensed mechanics.[16] Inspection of the software to confirm its method of controlling the tractors' functioning is technically illegal.

*Corporations may seek to manipulate health evidence or consumer behaviour.* Coca Cola has been criticised for providing financial and logistical support to the Global Energy Balance Network, a research organisation that promotes the idea that a lack of exercise, rather than the consumption of a sugar-rich diet, is the primary cause of obesity.[17] Health insurers are looking to use IoT monitoring of clients to better manage risk and reduce payments.[18] Amazon's WiFi connected 'Dash' device facilitates the

reordering of household products from Amazon at the push of a button, and when added to washing machines, the functionality allows the automatic ordering of washing powder when more is needed. The user does not see the cost of the transaction in either case.[19]

## 3. Discussion

There are several criteria by which the success of the IoT in healthcare might be evaluated. Ideally, any use of the IoT should: support the provision of effective, high quality, evidence-based care; protect patient safety; ensure patient privacy; secure data from loss during communication; and be free from commercial bias or distortion. However, each patient home could already have IoT devices such as light globes, refrigerators and doorbells, with different vendors, standards, and security and privacy settings. One insecure or poorly configured device within a home IoT network could provide access to most (or all) other devices on the network. There is almost certainly no competent technical professional curating or supporting this collection, making it challenging to meet the criteria outlined above.

Using the Internet of Things in the provision of healthcare will inevitably reproduce some existing risks associated with the use of ICT, such as threats to security and privacy, and the loss or delay of data in transit. However, the exponential increase of endpoint devices and intermediate equipment magnifies the number of items at risk, and places many devices in a poorly controlled environment, leading to changes in the likelihood of those risks. Use of the IoT also introduces new risks to patients from the exposure to commercial decisions without a mandated guarantee of ethical behaviour.

The German Advisory Council on Global Change (WBGU) has associated six clusters of systemic risk with characters from Greek mythology – Damocles, Cyclops, Pythia, Pandora, Cassandra and Medusa – the underlying narrative for each character matches the nature of the risk.[20] Damocles was invited to be 'king for a day'. On taking the throne, he found above him a sword hanging by a slender thread; 'Damocles' type risks have a high probability of damage, but a low likelihood. Pythia, a noted Delphic oracle, gave prophecies, which were highly valued, but ambiguous and hard to interpret; both the probability of 'Pythia' risks and the extent of damage are uncertain. When Pandora married, the gods gave her a closed box containing many evils, and told her to keep the lid closed. She opened it, releasing those evils upon the world. 'Pandora' risks involve extensive changes with no clear link to resulting damage; adverse effects emerge only after extensive spread of the changes has occurred.

The evolving risks associated with the IoT are unlikely to be risks of Damocles. They are more likely to have the nature of Pythia – neither the evidence of risk nor the extent of the damage will be immediately apparent. Over time, those risks will exert their effect, and in retrospect an absence of prior caution will become apparent. Risks characterized by Pandora may also eventuate. The changes brought by the IoT for healthcare will result in damage in the future, but the connection between the risk and the damage will not be immediately obvious. Once the link is made, it will be too late to reverse the changes. Renn and Klinke [21] identify the precautionary principle as the appropriate strategy for managing both types of risk.

What does the precautionary principle entail for a health IoT? As an ideal, new implementation should be supported by sound evidence of patient outcomes, safety, and benefits, with those benefits outweighing the costs of implementation, use and maintenance. There should also be a well-defined framework for the long-term

management of IoT devices, including the issuing and revoking of trust certificates, updating software, and eventual device retirement. One way of ensuring these ideals are met would be through control and regulation, although it seems that the time for such an approach may have passed, with no regulations in place. While most providers of health ICTs are ethical, focus on the needs of patients, and would comply with regulation, we should expect there would be a handful of individuals and firms who do not. Regrettably, this brief overview highlighting some of the likely risks inherent in health IoT cannot provide a clear way forward, but should serve as a warning that the precautionary principle must be applied.

## References

[1] ISO/IEC JTC 1. Internet of Things (IoT) Preliminary Report 2014. 2014.
[2] Smith SW, Erickson JS. Never Mind Pearl Harbor–What about a Cyber Love Canal? IEEE Secur. Priv. 2015 Mar;13(2):94–8.
[3] OECD, editor. Emerging systemic risks in the 21st century: an agenda for action. Paris: OECD; 2003.
[4] Kozlov D, Veijalainen J, Ali Y. Security and Privacy Threats in IoT Architectures. ACM; 2012 [cited 2015 Oct 6]. Available from: http://eudl.eu/doi/10.4108/icst.bodynets.2012.250550
[5] Vidalis S, Angelopoulou O. Assessing identity theft in the Internet of Things. IT Converg. Pract. 2014;2(1).
[6] Reback S, Costello T. Deconstructing the Internet of Things. Bloom. Gov. [Internet]. 2014 [cited 2015 Oct 6]; Available from: http://www.webcitation.org/6dvhzdKPn
[7] Papoutsi C, Reed JE, Marston C, Lewis R, Majeed A, Bell D. Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: results from a mixed methods study. BMC Med. Inform. Decis. Mak. [Internet]. 2015 Dec [cited 2015 Oct 23];15(1). Available from: http://www.biomedcentral.com/1472-6947/15/86
[8] Hill K. The Half-Baked Security Of Our "Internet Of Things" [Internet]. Forbes. 2014 [cited 2015 Oct 22]. Available from: http://www.webcitation.org/6dviHRyxN
[9] Whittaker Z. New security flaws found in popular IoT baby monitors [Internet]. ZDNet. 2015 [cited 2015 Oct 22]. Available from: http://www.webcitation.org/6dviMrwO4
[10] Pauli D. Thousands of "directly hackable" hospital devices exposed online [Internet]. The Register. 2015 [cited 2015 Oct 22]. Available from: http://www.webcitation.org/6cTqrpe1w
[11] Center for Devices and Radiological Health. Safety Communications - Cybersecurity Vulnerabilities of Hospira Symbiq Infusion System: FDA Safety Communication. 2015.
[12] Gupta S. Dick Cheney's heart [Internet]. CBS News. 2013 [cited 2015 Oct 22]. Available from: http://www.cbsnews.com/news/dick-cheneys-heart/
[13] Miller ME. How hackers can control your car from miles away [Internet]. Wash. Post. 2015 [cited 2015 Oct 22]. Available from: http://www.webcitation.org/6dviRi3xR
[14] Topham G, Clarke S, Levett C, Scruton P, Fidler M. The Volkswagen emissions scandal explained. The Guardian [Internet]. 2015 Sep 24 [cited 2015 Oct 23]; Available from: http://www.webcitation.org/6dviXUjcl
[15] Neslen A. Samsung TVs appear less energy efficient in real life than in tests | Environment | The Guardian [Internet]. 2015 [cited 2015 Oct 5]. Available from: http://www.webcitation.org/6dvibkrx2
[16] Sydell L. DIY Tractor Repair Runs Afoul Of Copyright Law [Internet]. NPR.org. 2015 [cited 2015 Oct 23]. Available from: http://www.webcitation.org/6dvifO3qx
[17] O'Connor A. Coca-Cola Funds Scientists Who Shift Blame for Obesity Away From Bad Diets [Internet]. Well - NYT. 2015 [cited 2015 Oct 23]. Available from: http://www.webcitation.org/6dvik99SZ
[18] Borst M. Insurance In The Age Of The Internet Of Things [Internet]. Digit. Mag. 2015 [cited 2015 Oct 5]. Available from: http://www.digitalistmag.com/industries/insurance-age-internet-things-02361459
[19] Fleishman G. Don't dash to Dash: new Amazon buttons aid brands, not consumers [Internet]. TechHive. 2015 [cited 2015 Oct 23]. Available from: http://www.webcitation.org/6dvipDXFB
[20] German Advisory Council on Global Change (WBGU), editor. World in transition: strategies for managing global environmental risks. Berlin: Springer; 2000.
[21] Renn O, Klinke A. Systemic risks: a new challenge for risk management. EMBO Rep. 2004;5(1S):S41–6.

# A New Challenge to Research Ethics: Patients-Led Research (PLR) and the Role of Internet Based Social Networks

Eugenia LAMAS [a,1], Rodrigo SALINAS[b], and Dominique VUILLAUME[c]

*aINSERM, U 1142, LIMICS, F-75006, Paris, France; Sorbonne Universités, UPMC UniversitéParis 06, Université Paris 13, Sorbonne Paris Cité, F-93430, Villetaneuse, France.*
*bFacultad de Medicina, Universidad de Chile, Santiago, Chile.*
*cCERMES 3 : INSERM U.988/UMR CNRS 8211/EHESS/Université Paris Descartes, F-94801, Villejuif, France.*

**Abstract.** A characteristic feature of the development of health-related social networks is the emergence of internet-based virtual communities, composed of patients. These communities go beyond the mere interchange of information concerning their conditions, intervening in the planning and execution of clinical research, including randomised controlled trials, in collaboration with health professionals. That was the case, in 2009, when patients suffering amyotrophic lateral sclerosis, a rare and severe disease, conducted a clinical trial in USA, organising themselves through an online platform. This initiative launched a new model for the planning and conduction of clinical research: "Participants-Led Research" (PLR). The distinctive particularities of this new research paradigm represent a challenge to the traditional standards used for judging the ethical soundness of clinical investigation. That is the case, for example, of informed consent. This article aims at identifying the ethical, legal, and social issues (ELSI) posed by PLR and the relevant concepts that may help in solving them. The following issues, in particular, are analysed, that may give place to a new social contract for the ethical assessment of clinical research: consent for participating in research and personal integrity; data protection and confidentiality; benefits sharing and intellectual property

**Keywords.** Community-Based Participatory Research, Ethics Research, Social Networks, Data

## Introduction

The development of health related social networks (health devoted websites and social networks, discussion forums, blogs, and patient's information and support websites, etc.) helped the creation of a novel kind of patient communities: Online Patient Networks, whose members interact among themselves through a number of different Internet based social networks. Some of these networks gather together patients suffering the same disease, usually severe and chronic.

---

[1] Corresponding Author.

The members of these communities take advantage of these networks to interchange disease related real-world experiences, current treatments and, in particular, available drugs, their adverse effects, recent scientific research results and ongoing randomised trials, concerning their conditions.

There is a number of examples, in the United States of America, of such communities going beyond the mere interchange of relevant information, involving themselves in either observational or experimental clinical research, together with health professionals, sharing information in a peer to peer manner, in order to generate relevant new data for their conditions. A well-known example of this sort of cooperation is the online platform PatientsLikeMe, which organised, in 2009, the lithium study for amyotrophic lateral sclerosis (ALS), a severe and rare neurological condition, launching a new paradigm for planning and conducting clinical research [1-2].

Since 2012, a number of papers have coined the expression "crowdsourced health research studies" (CHRS) to designate this new health research model, pointing out to one of its original features, that of mobilising multiple sources of information to help health research, taking advantage of the dynamic group interactions existing in virtual communities and the communication tools currently available in internet. Melanie Swan distinguishes two different types of CHRS: one of them organised by clinical researchers (research organised CHRS), and the other led by patients themselves (participants-organised CHRS) [3]. The lithium study for ALS is an example of the latter.

Other authors have proposed the expression "apomediated research" to refer to this sort of investigations, pointing out to the fact that this type of research finds its way through the direct exchange of information and data among the members of the virtual communities, avoiding the traditional collection of data, mediated by specialised health professionals: physicians, pharmacists and clinical researchers [4]. Günther Eysenbach had used the voice "apomediation" in his framework describing Medicine 2.0 [5].

Recently, stakeholders of this new model of investigation have proposed the term "participants-led research" (PLR) to describe it, emphasising its participative nature, identified as one of its original and characteristic features [6]. As a matter of fact, all these expressions describe this novel research model, highlighting different distinctive aspects of its innovative structure.

Although recent, this new model of research has already motivated some publications (au lieu de "a significant number of"). One of them evaluates the potential value of this innovative approach, comparing it with the traditional one. Using this plan, the ethical, legal, and social issues (ELS) are examined. It has been, thus, proposed, that a new ethical and legal regulatory framework is needed to address the particularities of CHRS, different from that currently in use to examine standard research, nonetheless serving the same purposes [7].

After defining PLR, a recent contribution examines the potential benefits and risks associated to the growing of this research strategy. The authors recognise that the scientific value and social utility of PLR requires a new social contract, giving a thorough account of its relevant features [6].

Indeed, the PLR has so many particularities embedded in its structure, that they challenge the pertinence of the ethical principles themselves, which support the current ethical evaluation system. This is the case, for example, of informed consent, a key component of ethically sound research, whose importance could be put to question, as part of a research model in which those planning and conducting the investigation, act

as research subjects themselves [8]. This may result, for example, in a challenge to Research Ethical Committees (RECs) role, when evaluating these peculiar informed consents, aimed at protecting research participants from themselves.

In this context, this report aims at reflecting on some ELS issues concerning PLR. In particular, the protection of personal integrity, the confidentiality of data, and the new paradigm for data sharing and intellectual property in PLR generated information.

## 1. Methods

Our prospective analysis is based on a comprehensive research of published reports, indexed in relevant databases, addressing research led by patients' communities, through digital networks. In order to narrow our search, increasing the specificity of its results, we aimed at collecting only those reports addressing participants-organised CHRS, adopting the classification proposed by Swan [3]. We also considered those reports examining "research-organised" CHRS, in which professional researchers lead the initiative, to provide a comprehensive background to our analysis.

## 2. Results and Discussion

We identified three major new issues related to PLR and the emergence of patient's virtual community's involvement in research.

PLR shifts the equilibrium of power, from professional researchers to research subjects, differentiating this innovative model from traditional research. In PLR, patients not only participate in organising and conducting the investigation, but also on the recollection and analysis of the data resulting from the research, vis-à-vis those professionals participating in the investigation. As a result of this involvement, the centre of gravity of research is displaced from the professional interests to those belonging to the patients' community, thanks to their active involvement through digital networking. This collaboration permits a previously inexistent mix of lay expertise (coming from those patients acting as research subjects themselves) and professional expertise (coming from those physicians and pharmacists that collaborate with the execution of the study and the analysis of the data).

This entirely new situation, blurring the border between the patient as passive participant and its new role as protagonist of the research, actively participating in planning and conducting the investigation, generates a number of previously unsuspected issues, concerning the notion of personal protection. In fact, this notion has to be understood not only as protection against research related risks, but also as protection and confidentiality of personal health information, in a setting in which patients are both actors and participants in clinical research.

In PLR, patients provide their personal data, voluntarily sharing them with other patients, in order to get useful information for the diagnosis and treatment of their condition. When sharing the data, they resign their right to the confidentiality of this information, overcoming this traditional principle of human subject's research, as a demonstration of "self-empowerment" of patients themselves. A novel challenge to the role of RECs is to discern if they are entitled to protect patients from themselves, concerning their data protection, as they were originally meant to be protected from third parties, engaged in human subject's research.

This dilemma might be considered as a result of the progressive differentiation between the personal and public spheres, originating in the eighteenth century enlightened western society, empowering the individual with, nowadays, universally recognised rights. These rights, aimed at protecting the individuals from the intrusion of the state and the dominant class in their private sphere and decisions, were not meant to protect individuals from their own decisions, or to compel them to behave in a certain manner to avoid being harmed. The role of RECs in the protection of human subjects is consistent with this enlightened approach, and the recognition of their rights is enshrined in a number of covenants and declarations. PLR, as a new paradigm of patients' empowerment in research, challenges this traditional approach, questioning the meaning, for example, of data protection in this new setting and the role of RECs in overseeing the fulfilment of this right.

Is it conceivable that human subjects might resign the right to protection of their data, when participating in PLR? It is widely accepted that some rights are unwaivable, for example the right to not be enslaved, but some authors pose that this is not the case with every human right: "…one cannot waive one's rights to autonomy and liberty, [but] one probably can, in certain circumstances, waive one's human right to privacy" [9]. It is debatable if this is the case with data collected in PLR, but it is beyond doubt that this is a major issue that has to be addressed, when discussing the appropriateness of current RECs regulation. Some authors have pointed out, even, the need of a new social contract for fulfilling the functions of these committees [6].

The second major issue we identified, as a result of the blurring border between the patient and the researcher, is related with the purpose and the quality of the data obtained from PLR. This new kind of research, concerning the knowledge and therapeutics of certain diseases, represent a valuable source of information in certain areas, which not always is easily obtained from traditional research, as it is the case of pharmacovigilance, for example. This new way of collecting and assembling data, however, compels the scientific community to check the quality of the information that is gathered by PLR, and the need for empowering this new sort of citizens' generated science with the required tools for ensuring the quality and reproducibility of the data. A major challenge to PLR, thus, is represented by the need of professionalizing the role of patient-researcher, when feeding the results of clinical research with self-reported data, to match the current standards of traditional investigations proceedings, and avoiding the potential multiplication of error occasioned by a multi-source of data collection. This massive involvement of patient-researchers, at the same time, represents a precious opportunity to improve the statistical power of clinical research, and the common difficulties in finding and enrolling patients. Vayena et al. [6] consider this issue, concerning the quality of data generated by PLR, as a key part to be considered in the new social contract that needs to be agreed by the scientific and lay communities, when taking charge of these new developments.

A third issue, singled out in our search, and not previously identified by those who have systematically addressed this subject, is posed by the need of discussing new strategies regarding intellectual property, and the sharing of benefits originating in this new kind of research, led by patients, and with the occasional collaboration of academic and for-profit sectors. This discussion is particularly relevant for those chronic and rare conditions, in which the collaboration of patients' organizations is crucial for the success of clinical research [10].

In summary, PLR represents, at the same time, an innovative manner for organising clinical research, a precious opportunity for improving pharmacovigilance

[11], for increasing the enrolment of patients, particularly in uncommon conditions, and –most notably- a challenge to the framework used by RECs for the standard ethical assessment of investigation on human subjects.

## References

[1] P. Wicks, T.E. Vaughan, M.P. Massagli, J. Heywood, Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm, *Nature Biotechnology,* **29** (2011), 411-414.

[2] P. Mansell, PatientsLikeMe lithium study shows real-time value of social networks, *Clinical News,* April 26 (2011), 30-31

[3] M. Swan, Crowdsourced Health Research Studies: an Important Emerging Complement to Clinical Trials in the Public Health Research Ecosystem, *J. Med. Internet Res.***14**(2) (2012), e46

[4] D. O'Connor, The Apomediated World: Regulating Research When Social Media Has Changed Research, *Journal of Law, Medicine & Ethics,* **41**(2) (2013), 470- 483.

[5] G. Eysenbach, Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness, *J. Med. Internet Res.,* **10**(3) (2008), e22.

[6] E. Vayena, R. Brownsword, S.J. Edwards, B. Greshake, J.P. Kahn, N. Ladher, J. Montgomery, D. O'Connor, O. O'Neill, M.P. Richards, A. Rid, M. Sheehan, P. Wicks, J. Tasioulas, Research led by participants: a new social contract for a new kind of research, *J. Med. Ethics,* doi: **10. 1136**/medethics (2015), 102663

[7] M. Swan, Scaling crowdsourced health studies: the emergence of a new form of contract research organization, *Personalized Medicine,* **9**(2) (2012), 223-234.

[8] E. Check Hayden, A broken contract, *Nature,* **486** (2012), 312-314.

[9] J. Griffin, *On human rights,* Oxford University Press, Oxford, 2008.

[10] R. Giffin, S. Robinson, T. Wizemann, Breakthrough Business Models: Drug Development for Rare and Neglected Diseases and Individualized Therapies: Workshop summary, National Academies Press, Washington D.C., 2009.

[11] J. Lardon, R. Abdellaoui, F. Bellet, H. Asfari, J. Souvignet, N. Texier, MC.Jaulent, MN. Beyens, A. Burgun, C. Bousquet. Adverse Drug Reaction Identification and Extraction in Social Media: A Scoping Review. *J Med Internet Res.* **17**(7) (2015), e171

# Analyzing Privacy Risks of mHealth Applications

Alexander MENSE [a,1], Sabrina STEGER [a], Matthias SULEK [a], Dragan JUKIC-SUNARIC [a], and András MÉSZÁROS [a]

[a] *University of Applied Sciences Technikum Wien, Vienna, Austria*

**Abstract.** Mobile health applications are expected to play a major role for the management of personal health in the future. For this purpose, the apps collect a lot of sensitive data from sensors or direct user input, combine it with automatic data such as GPS location data, store it locally and pass it on to web-platforms (often running in a public cloud), where the information can be managed and often shared with others in social networks. However, it is usually not transparent for the user how this sensitive information is handled and where it goes to. This paper shows the result of the analysis of mobile health applications regarding the handling of sensitive data especially with respect to transmission to third-parties.

**Keywords.** Mobile Health, Security, Privacy

## Introduction

Smartphones, tablets, and mobile applications/apps are taking over our daily lives more and more. People are getting more and more addicted to their smartphones. With the help of modern features such as GPS tracking, the smartphone stores a lot of information about us, e.g. where we live, where we sleep, where we work, etc. Almost every app collects a certain amount of personal information about the smartphone user.

Among millions of apps, there are thousands of different mHealth apps, which are supposed to help us lose weight, track our fitness level, or create diet plans. Many of these apps seem to be every useful as personal health assistant, but they actually handle a lot of sensitive and private information that requires appropriate secure handling to protect privacy.

Wu-Chen Su [1] reviewed existing relevant research about smartphone applications in the eHealth domain and identified a set of pertinent challenges. One of these is about security and privacy concerns, which have to be further explored and discussed by researchers. This paper presents the results of the technical analysis of mHealth applications running on Android regarding privacy and security risks.

## 1. Methods

The evaluation process involved reviewing mHealth applications by primarily analysing data sent over a network communication between the applications and the

---

[1] Corresponding Author.

Internet. A set of selected ten free applications out of several hundreds available in the "health and fitness" category was systematically selected to represent a broad range of personal health functionalities: personal health record, self-management, calorie counter and diet plan, healthy living and health promotion (activity and fitness tracker, workout, and sports), and medication management.

To provide uniform parameters for every application, a specific test environment made of one specific device was used. Apps were tested on an Android emulator based on a VirtualBox called Genymotion (https://www.genymotion.com). It can emulate specific devices, thus making it a good choice to test mobile applications. The test environment was set up for a Google Nexus 4 device with Android 4.3. To get access and for testing purposes, a Google test account has been created.

Examination and analysis of the data traffic between mobile applications and the internet was done with a proxy which worked like a "man in the middle attack" (operated on a laptop with Windows OS). Furthermore, the proxy was also able to intercept SSL encrypted traffic. The proxy therefore dynamically creates a certificate for the server and signs it with its own root certificate, which is needed to be installed as a trusted certificate on the mobile device. With this configuration, it is possible to debug secure (SSL) communication as well. The test environment is shown in Figure 1.
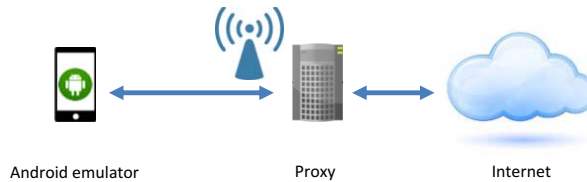


**Figure 1**. Test environment.

Mobile applications were used according to the intended standard use. They were installed from Google play store, necessary registration and login was done with a test user, and all offered functions were used.

Data collected by the mobile applications were analysed including user input (e.g. name, gender, health data …) and background information (e.g. GPS, device identification, contact data …). Furthermore, it was investigated whether the applications encrypt the traffic and with how many different domains they exchange information.

## 2. Results

In a first step, the permissions required by the applications were examined. In almost all cases, users have to grant a broad range of permissions including for instance in-app purchases, access to identity information, contacts, location, photos/media/files, camera, microphone, Wi-Fi-connection information, and Bluetooth connection information.

In a second step the network traffic was analysed. 90% of the evaluated mobile applications communicate with the application developer-controlled website and/or with third-party domains. In fact, only one application did not transmit data at all, but stores it in an unencrypted csv-file located on the SDcard which can be easily accessed and read by any other application or malware.

## 2.1. Communication with the developer-controlled website

Last year, Symantec [2] published that 20% of a set of analysed tracking applications (not limited to health) transmitted user login credentials in clear text. In 2013, Lie Njie [5] analysed mobile health and fitness applications and discovered that only 15% of them encrypted communication with SSL to the developers' website. Furthermore, Lie Njie [5] found that none of the applications sent the data to third-party advertisers using an encrypted SSL connection. In contrast, in our setup, 100% of the evaluated applications use encrypted (SSL) communication with the developer-controlled website. If applications use the HTTPS protocol it makes it harder to intercept data – but not impossible. By installing a man in the middle certificate it was possible to intercept the traffic of all of the applications, which means none of the applications does appropriate certificate checks and certificate pinning.

## 2.2. Communication with third party (advertising, analytics) sites

Eighty percent of the analysed free mobile applications contact third-party websites for advertising and analytics. The communication with the advertising sites occurred mainly unencrypted. One of the analysed applications sends also "usage data" in plain text to third-party advertisers after e.g. measuring fitness activities (see figure 2).

```
os=android&model=genymotion google nexus 4 - 4.3
-    api    18    -    768x1280&keywords=Run,run,
jog&oauth_key=5482,5410,&os_ver=4.3&locale=en_US&did
=43bc618c5de25126d5309ec7bda26b2f&postal_code=8783&w
orkout_pace=0.0&age_group=25to34&workout_time=4&gend
er=female&carrier&user_id=64422724&workout_heartrate
=0&workout_calories=0&workout_distance=0.0&workout_s
peed=0.0&make=genymotion&app_ver=3.5.1&app_store=goo
gle.
```

**Figure 2.** Sending usage data.

Another application transmits the GPS location (latitude and longitude) unencrypted to a third-party advertising site.

On the other hand, it is not transparent to users how much additional data regarding the usage of the app and the mobile device is collected and sent to third-party sites for analytics purposes. This form of data collection is known as behavioural tracking [5]. Ninety percent of the analysed applications provide such information to different sites. While Symantec examined a self-tracking application which contacted 14 domains [2], in our tests the number varied between two and ten different domains. Many different URLs had been identified: admob.com, appsfyler.com, flurry.com, fiksu.com, google-analytics.com, localytics.com, kiip.me, rubiconproject.com, crashlytics.com, newrelic.com. One application sends the data even unencrypted to http://data.flurry.com, while all others use encryption also to transmit analytics data.
Ninety percent of the inspected apps transmit more information than what is necessary for the proper running of the app, e.g. users' location (GPS), android version, phone information etc. (figure 3).

```
{"app":{"versionCode":274,"app_key":"9fa34770423d
7c3c3f1e58a1620f5d49","version":"274
5.2","versionName":"5.2"},"connection":{"type":"WIFI
","carrier":""},"source":"application","location":{"
accuracy":40,"time":"2015-02-
17T06:03:13.000","lng":14.550071666666668,"lat":47.5
1619833333333}…  "device":{"id":"1290319f-ab68-4584-
9a31-53b5163d6869","os":"Android
4.3","kiip_uuid":"963a3679-4d59-4f76-a1d0-
a95598a013f7","timezone":"Europe\/Vienna",…}
```

**Figure 3.** Sending information to analytics sites.

It is obvious that applications contact remote servers for some of their functionalities (to get images and marketing advertisement), nevertheless the number of contacted third-party analytic websites is surprisingly high.

## 2.3. Device information

The IMEI (International Mobile Equipment Identity) is supposed to be a unique identifier for a device and can never be changed. In 2010, the Wall Street Journal reported that they had analysed 101 popular smartphone applications and had identified 56 applications transmitting the unique device ID. This occurred without users' awareness [4].  The analysis of the encrypted requests shows that 30% of the applications send the device ID to the application developers' websites. For example, to track the user's sport activities, a fitness app delivered the device ID (not IMEI) with each request. While the user enters his/her weight, the device information and the device ID are transmitted. The information sent is shown in figure 4 below:

```
deviceApp=paid%2C5.2&device=android%2Cvbox86p%2C4
.3%2C18%2CGoogle+Nexus+4+-+4.3+-+API+18+-
+768x1280&weight=55.0000&dateStr=1423687845633&email
=susi123.mustermann%40gmail.com&apiVer=2.3&deviceID=
ed380bbc-859d-4573-9caa-49800678677c.
```

**Figure 4**. Usage of the unique device identifier.

## 2.4.  Contact information

In addition to the device information, applications were suspected to collect information from the local contact list. We could prove that e-mail addresses but not phone numbers were collected and delivered encrypted to the application developers' website. For example, a fitness application tracking running activities delivers the contact e-mail addresses to help the user find his or her friends faster (figure 5).

```
emails=%5B%22Sandra.Huber%40gmx.at
%22%2C%22Nicole.Hahn%40gmx.at%22%5D&fbuids=%5B%5D
```

**Figure 5.** Usage of contact information.

## 3. Discussion

An issue to be discussed is the location of the servers where data is transmitted to. While the analysis shows that 90% of analytics and advertising servers are located

outside Europe, the location of the developer-controlled servers highly depends on the country the developing company resides. Sending sensitive data to outside of the EU is bound to privacy law restrictions, which are usually handled by letting the user explicitly agreeing to the terms & conditions, resp. the privacy policy.

A recently published study from Huckvale et al [10] about the analysis of apps included in NHS England's Health Apps Library, that actually shows results similar to those described in our analysis (expressing the lack of security and privacy of many mHealth apps) evaluated privacy policies. They show that most apps do not handle data according to their privacy policy, and that some apps do not even have any privacy policy. Besides, it's well known that such policies are often quite large and complex documents that users mostly care as low as they care for the rights an app asks for.

Mitigation of security and privacy concerns of mHealth apps has come in the focus of the European Commission and is addressed in the "Green Paper on mobile Health (mHealth)" [11] from technical up to legal levels.

Technical measures (independent of the OS and apps) to offer users a minimum of control over the transmission of sensitive data are still quite limited. A first technical approach would be the use of a privacy proxy that blocks unwanted traffic. It may analyse data streams to determine unwanted traffic to ad-sites as well as analytics-sites and can block these connections. A second approach could be the active filtering of sensitive information or the provision of fake information. Sensitive information is removed from the unwanted communication streams by using for instance taint analysis techniques to analyse data flows and mark (taint) sensitive values. Also fake information could be delivered to tracking and advertising networks. Users can define to send random or an explicit amount of fake answers (e.g. PDroid-Tools).

## References

[1] Wu-Chen Su, A Preliminary Survey of Knowledge Discovery on Smartphone Applications (apps): Principles, Techniques and Research Directions for E-health, International Conference on Complex Medical Engineering, pp.369-374, IEEE Computer Society Press, June 2014.

[2] M. B. Barcena, C. Wuesst, H. Lau, "How safe is your quantified self?" Symantec, August 2014.

[3] N. Elenkov, "Android Security Internals: An In-Depth Guide to Android's Security Architecture," No Starch Press, October 2014.

[4] S. Thurm, Y. I. KANE, "Your Apps Are Watching You," The Wall Street Journal, December 2010.

[5] L. Njie, "Technical Analysis of the Data Practices and Privacy Risks of 43 Popular Mobile Health and Fitness Applications," Research Performed For: Privacy Rights Clearinghouse, July 2013.

[6] Google Android Dev, "http://developer.android.com/guide/topics/admin/device-admin.html," Device Admin, Last Access: 18th February 2015.

[7] Google Android Dev, "http://developer.android.com/reference/ android/app/admin/DevicePolicyManager.html," Device Policy Manager, Last Access: 18th February 2015.

[8] A. B. Jeng, L. Hahn-Mingm, C. Chen, C. Tien, „Android Privacy," IEEE, Dept. Of Comput. Sci. & Inf. Eng., July 2012.

[9] Forbes, "f you're not paying for it, you become the product", http://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-become-the-product/", Last Access: 18th February 2015.

[10] Kit Huckvale, José Tomás Prieto, Myra Tilney, Pierre-Jean Benghozi, and Josip Car, Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Medicine 2015, 13:214. doi:10.1186/s12916-015-0444-y

[11] European Commission, Green Paper on mobile Health ("mHealth"). http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=5147, Last Access: 6th January 2016

# Non-technical Issues in Design and Development of Personal Portable Devices

Lenka LHOTSKA [a,1], Paul CHESHIRE [b], Peter PHAROW [c], and David MACKU [a]

*[a] Czech Technical University in Prague, Czech Republic*
*[b] Independent Consultant, Kent, United Kingdom*
*[c] Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany*

**Abstract.** Mobile technologies are constantly evolving and with the development of Internet of Things we can expect continuous increase of various applications. Mobile technologies have undeniable opportunities to play an important role in health services. Concerning purely technical aspects, almost every problem can be solved. However, there are still many unsolved and unclear issues related with ethics and governance mechanisms for mobile phone applications. These issues are even more critical in medical and health care applications of mobile technologies. This paper tries to analyse ethical, and privacy-related challenges that may occur when introducing Personal Portable Devices (PPD) to collect and record personal health data in health care and welfare environment.

**Keywords.** Personal Portable Devices, Ethics, Human Factors

## Introduction

The integration of mobile technologies, sensors, sensor networks and portable devices to provide personal services of many kinds, especially those involving location tracking or lifestyle and well-being, will require wide acceptance of these technologies by ordinary citizens. When we design and develop applications for health care and welfare, we have to consider users related issues much more carefully than in other more general applications. Recently, at many conferences and in many journal articles, various technological issues concerning Personal Portable Devices (PPDs) have been discussed and many new solutions have been proposed. With the advent of Internet of Things and their applications in health and social care, the requirement of wide acceptance is becoming more urgent. However, the non-technical dimensions of the applications remain almost unnoticed. Whilst it is clear that many portable technologies are widely used by people of all ages and abilities, their field of application is more "recreational" than "clinical". Without giving adequate consideration to psychological, ethical and legal issues it would be very difficult to introduce any technological solution that supports or replaces healthcare professionals into routine use. In this paper we try to identify and discuss the most important problems that should be solved (at least to some extent) before portable "clinical" technology is planned to be introduced to the market.

Since it is assumed that these systems will be mostly used by people who are not technology specialists (care givers, patients, their families, social workers, etc.) it is

---

[1] Corresponding Author.

necessary to focus even more on: the acceptability of these devices by everyone in the chain of care, human-computer interaction, easy and intuitive control, prevention of misuse, etc. In addition to these items, there are many questions linked with data collecting, storing, processing, access, and use. We try to analyse the major legal and ethical issues which arise from the handling of sensitive data about health and data about everyday activity patterns. Since the space is limited we focus on ethical issues in relation to different modes of use of personal portable devices and mobile applications.

## 1. Methods

In our research we decided to analyse the management of ethical issues connected with mobile technologies and PPDs in different applications and different settings. We tried to report quantitative information about the growth of number of users and applications advocating the importance of development of a certain regulatory framework. When relevant, we mention existing legal regulations or guidelines. It is necessary to distinguish for which purpose the mobile technologies and PPDs are used, who interprets the acquired data and how the patient is informed about his/her state. From the regulatory point of view there is rather strict division of applications for medical purposes on one side and fitness and well-being on the other side.

Mobile technologies evolve constantly and with the advance of Internet of Things we can expect continuous increase of applications installed in persons´ homes, senior homes as well as an increase use of wearable's for measurement of vital parameters and behaviour monitoring. Since 2007 we observe continuous growth of mobile devices. Number of mobile phone users has increased from 400 mil. users in 2007 up to almost 2 bil. in 2015. It is expected that by 2017 the global coverage by commercial wireless signal will reach 85% of the world´s population [1] and approximately 3.4 bil. people worldwide will own a smartphone. Half of them will use health apps.

However, there are still many unsolved and unclear issues concerning ethics and governance mechanisms for virtual technology and for mobile phone applications. In this respect, a significant change happened in September 2013, when the FDA released guidance for the developers of mobile medical apps [2]. „Mobile medical apps" are defined as applications which are intended „to be used as an accessory to a regulated medical device" or „to transform a mobile platform into as regulated medical device ".

In the EU, there are no binding rules as to the delimitation between lifestyle and wellbeing apps and a medical device or in vitro diagnostic medical device. As the use of these apps is affected by existing EU regulatory instruments, stakeholders, such as mobile app developers and mobile platform manufacturers, may seek guidance as to the applicable rules. Since January 2012, in order to help software developers and manufacturers to identify whether their products fall or not under the Medical Devices Directive (Directive 93/42/EEC on medical devices, OJ L169,12.07.1993) or the in vitro diagnostic medical devices Directive (Directive 98/79/EC on in vitro diagnostic medical devices, OJ L331, 7.02.1998) [3], the Commission's services have issued some guidance on this issue, which will be continuously updated. According to this guidance, depending on their intended purpose, apps may fall under the definitions of a medical device or of an in vitro diagnostic medical device and consequently will have to comply with the relevant provisions of the aforementioned directives. Since this delimitation is not yet clarified through binding rules, when the Medical Devices Directives do not apply to apps, clarity is required as to the rules with which they must comply. The fact, that Union

legislation could not yet address latest developments in this sector and that the Court has not had the opportunity to clarify the applicability of existing legislation on these newly developed apps, still leaves room for interpretation. Thus in April 2014, the European Commission launched a public consultation alongside the Green Paper on mobile health (mHealth) [4] to help identify the right way forward to unlock the potential of mobile health in the EU. Together with the Green Paper, the Commission also published a Staff Working Document on the existing EU legal framework applicable to lifestyle and wellbeing apps [5], providing legal guidance on EU legislation in the field to app developers, medical device manufacturers, digital distribution platforms, etc.

In the world there are over 150 countries that have to develop any kind of regulatory framework for apps. It will be a complex process because it is impossible to apply one model of mHealth solution to all these countries. The reason is that every country has a unique health system with its own legal regulations, which these apps must fit with. However, even existence of guidelines and a framework does not guarantee quality of the apps as findings in [6] illustrate. The study published in the New England Journal of Medicine in 2014 shows that in the USA only about 100 of 100 000 health care apps are FDA approved. Many of the apps are used for collecting and analysing data from basic wearable's, whose market is also continuously growing. In [7] FierceMarket informs that while in 2014 26.4 mil. basic wearable's (such as fitness wristbands) were shipped, in 2015 the number significantly increased to 72.1 mil. Therefore, the importance of validation of the apps becomes more pronounced.

## 2. Results

Based on the analysis and our practical experience we identified several groups of applications that differ by their purposes. The first group aimed at clinical research uses the collected data in controlled manner might be purely software app or combination of a medical device and software. The data are interpreted and verified by medical experts. The second group utilizing purely mobile applications developed in frame of clinical research and transferred to routine use collects data from the patients. The data are sent for interpretation to a medical centre. The third group, aimed at general population, can be either mobile apps in smart phones or applications collecting data from, for example, fitness wristband. Regarding output of the applications, the first two groups might represent decision support systems, while the last group represents recommender system [8]. In all cases the data are transmitted: from device to mobile phone and then to a medical centre. Although data privacy and security are well defined and the technological tools are well developed there are many problems and insufficiencies in existing applications. Recent study [9] revealed failures even in certified apps. Another issue which we did not find in any source is data persistence in mobile apps. For long-term monitoring of health state, we want to have access to past data. What happens when the user changes the type of the smart phone, brand or operating system? Is there any chance to transfer the stored data in format readable on the new device?

The major ethical issues arise from the handling of sensitive data about health and daily activities. Explicit informed consent must be obtained from the participants in order to include this sensitive data in their Electronic Health Record (EHR) or on local storage. They must also consent to their data being shared, transmitted and analysed by authorized personnel within the designed system. In mobile apps it is sometimes not clear who might have access to the data or where the data are sent.

All the procedures must conform to relevant EU legislation (in EU countries) and to national legislations related to the principle of respecting confidentiality. Let us consider a case of cross-border utilization of the system and data communication between two EU countries. The applications must be compliant with legal frameworks of both countries. Even in EU the situation is not that simple. Although it exists EU directives, the national legal regulations usually have some additional restrictions or requirements. A good example is data privacy. For example: Czech law does not explicitly state where the data should be physically stored. The only requirement is that the person/institution that collects such data ensures technical and organizational measures for data security and privacy. Spanish law puts additional requirements on the technical side and defines where and under which measures the data should be stored.

When designing any system that collects patient data all subjects must participate voluntarily after being informed of the objectives and methodologies of the project. Explicit written consent will be requested. Since some difference in respecting privacy could arise between different subjects, two different forms for informed consent / authorization should be used: 1) subjects who will authorize use of personal data; 2) physicians who consent to use their information and knowledge about patients who already agreed to participate. To be enrolled, both consents are needed.

All the researchers involved who take part in the analysis of non-anonymous data must be asked for an explicit declaration of respecting confidentiality. Personal data must not be used for commercial purposes. In the management of sensitive data, it is necessary to define proper levels of security. In most cases it is essential to assure both secure transmissions and secure storage and management of data and knowledge. Concerning development of an application, the following questions must be answered: Will be the behaviour of the patient recorded? Who will have access to the data? How long should the data be stored? Will the data be stored locally or in a central data storage? Are there any national legal regulations concerning the domain? How can the legal regulations influence the technological design and use of the proposed system?

Legal clarity, or rather the lack thereof, is a key issue. The lack of a clearly set out legal framework, in particular with regard to licencing, accreditation and registration of telemedicine, assistive technologies and ambient assisted living services and professionals, liability, reimbursement, jurisdiction, is a major challenge.

## 3. Discussion

Activities in medical areas are well reflected in ethical and legal documents. Thus the part of AAL handling data about persons' health state can be approached in the same way. However, the daily activities monitoring and other activities must be described appropriately with respect to ethical and legal frameworks valid in corresponding countries. Recently we have identified certain gaps in legal regulations on the border between health and social care in some countries with regard to application of mobile technologies, assistive technologies and currently Internet of Things (IoT). The critical issue as we see it from discussions at national level is data privacy, in particular who can have access to which part of patient/client data and how detailed information should be stored locally or globally and how long.

As we have already mentioned there are differences among EU countries in legal regulations. In addition, there are also differences in organization of health and social care. For example, in Austria both types of care are under the umbrella of one ministry;

in the Czech Republic health care is under the responsibility of the Ministry of Health and social care is controlled by the Ministry of Labour and Social Affairs. Thus the processes describing care provision are definitely different. This issue must be considered during the design and development of the applications. It will also influence processes necessary for cross-border provision of health care. Obviously in the future utilization of mobile apps across borders will be demanded more and more. However, the procedures are not clearly defined yet. Benefits of these apps will depend on how they are treated in the social context of the countries in which they are used. Wireless networks transcend borders, but not every country has the same health system and health care policy. Easier operation will probably lead to a requirement of certain process interoperability definition.

## Acknowledgement

## References

[1]    Savitz E. Ericsson: 85% global 3G coverage by 2017; 50% for 4G. *Forbes*. 6 May 2012. http://www.forbes.com/sites/ericsavitz/2012/06/05/ericsson-sees-85-global-3g-wireless-coverage-by-2017-50-4g-coverage/ (accessed 12 October 2014).

[2]    U.S. Food and Drug Administration. *Mobile medical applications: guidance for industry and Food and Drug Administration staff.* U.S Department of Health and Human Services. 2013.http://www.fda.gov/downloads/MedicalDevices/…/UCM263366.pdf

[3]    EU law and other EU public documents (online) retrieved on 26-10-2015 from http://eur-lex.europa.eu/homepage.html

[4]    GREEN PAPER on mobile Health ("mHealth") Brussels, 10.4.2014 COM(2014) 219 final {SWD(2014) 135 final} (online) retrieved on 10-02-2016 from https://ec.europa.eu/digital-agenda/en/news/green-paper-mobile-health-mhealth

[5]    COMMISSION STAFF WORKING DOCUMENT on the existing EU legal framework applicable to lifestyle and wellbeing apps Accompanying the document GREEN PAPER on mobile Health ("mHealth") Brussels, 10.4.2014 SWD(2014) 135 final {COM(2014) 219 final} (online) retrieved on 10-02-2016 from https://ec.europa.eu/digital-agenda/en/news/commission-staff-working-document-existing-eu-legal-framework-applicable-lifestyle-and

[6]    Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. *N Engl J Med* 2014;371:372-9.

[7]    http://www.fiercemobilehealthcare.com/story/why-security-must-be-top-focus-mhealth-wearable-data-exchange-strategy/2015-06-27

[8]    Pharow, P. - Lhotská, L. - Cheshire, P.: Personal Portable Devices as Enablers for Advanced pHealth Decision Support and Decision Making Services. In *Proc of the 10th Int Conference on Wearable Micro and Nano Technologies for Personalized Health*. Amsterdam: IOS Press, 2013, p. 27-32.

[9]    K. Huckvale, JT Prieto, M Tilney, PJ Benghozi, and J Car. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Medicine* (2015) 13:214

# Towards Citizen-Centred Care: Interim Results from an E-Prescription Case

Riikka VUOKKO [a,1], Outi TIKKALA [b], and Päivi MÄKELÄ-BENGS [a]

[a] *Finnish Institute for Health and Welfare (THL), Helsinki, Finland*
[b] *Lappeenranta University of Technology, Lappeenranta, Finland*

**Abstract.** Medication data is a crucial part of patient data. Medication data is stored in a centralized archive, and made accessible to health care professionals and citizens. Re-usability of medication data requires it to be not only interoperable, but also complete and reliable. We evaluated e-prescription data stored in a national archive. The data consists of 596 patients with 76411 e-prescriptions. The interim results show the data to be complete when stored, whereas data reliability would require more user training.

**Keywords.** health records, e-prescription, national health services, case study.

## Introduction

Patient data and personal health data is recorded into various systems and with different user needs [1]. Besides patient data citizens record health data into various devices and services. At the same time the availability of better, interactive e-health services is needed to utilize potential, citizen generated data. At the moment, citizens have access to their own patient data in Finland. In this regard, medication data is crucial both to health care professionals planning and carrying out treatments and care as well as to citizens, for example in checking their medication and dosage information. Comprehensive and accessible medication data increases patient safety. The national e-health services enable e-prescription data access regardless of the health care organization where it was first recorded. The citizens have a portal for accessing their e-prescription data stored in the national archive. Citizens can also request a physician to renew a prescription via a portal that is a part of the national e-health information services.

Medication data is well structured and already utilizable in, for example, decision support systems (DSS). Physicians and pharmacists have various DSS-systems and integrated EHR system tools available. Similarly, citizens would need support tools to interpret their medication and medication interaction or other essential information. Moreover, up-to-date medication data would be more complete when citizens can enter additional medication information of other medicinal products they are using. Previously, we have conducted a systematic literature review of the structured patient data from a secondary user viewpoint [2, 3]. The review clearly demonstrated the requirements for patient data quality; these include completeness and compatibility of the data with data structures, and data consistency and reliability in regard to data

---

[1] Corresponding Author.

retrieval. For example, there is evidence that using International Classification of Diseases (ICD) codes can increase both the completeness and reliability of EHR data content [4-6]. Data utilization in different systems and with various interfaces requires interoperable data, including personal health record (PHR) requirements. Recently it has been suggested that the different uses of patient data should be considered when defining data completeness [7]. If different data needs are addressed, the completeness of EHR data can be measured in comparison to the intended use. That is not only task dependent but also user dependent, as for example physician and citizens have different information needs.

In 2015, we carried out the first part of a study evaluating e-prescription data quality in the National Prescription Centre. In this paper, we give preliminary answers to our research questions: what is the quality of e-prescription data and is the data complete and reliable for user needs and, further, for secondary use purposes and developing digital services for both health care professionals and citizens.

## 1. Methods

The study was based on a quantitative analysis of e-prescription data. The analysis material was obtained at spring 2015, and the second author of this paper conducted the analysis using statistical analysis software. The analysis results were refined by all three authors.

The study materials included e-prescription data between 2012 and 2014 from all the patients whose prescriptions had been updated in November 2014. As a control group, we had patients whose prescriptions were not updated in November 2014. The gathered data consists of 596 patients with 76411 e-prescriptions. The control data has 600 patients with 39327 e-prescriptions. At the time of data retrieval, all public services providers had already integrated the e-prescription services.

The anonymised data covers structured prescription data, for example document identifiers, organization unit identifiers, prescription name, ATC-code, purpose of use, substance strength, prescribed quantity, dosage, dose form, route of administration, container, package data, and so on – in a total of 48 data elements. Data completeness was measured using quantitative comparison between different message types. Data reliability was evaluated as accuracy in regard to data specifications, as no more accurate measurement is currently available for comparison.

## 2. Results

Completeness of the medication data was analysed with quantitative comparison between message types and mandatory data elements. The data structure implemented for e-prescription orders consists of Boolean values and coded values to a degree, but a lot of the data is still entered as free text by the ordering physician. Much of the data entered into e-prescription is mandatory. A Boolean value indicating the start of a new medication order was entered in 92 % of the cases. In 80 % of these cases the Boolean had a *false* value, indicating the prescription order was a continuation of patient's ongoing medication. However, the Boolean value for permanent medication was entered only in 62 % of the cases. Data covering permanent medication is being entered differently in various hospital districts, with data entries ranging approximately from

40 % to 1 % of all entries within a district. Similarly, a Boolean indicating that a prescription order concerns a narcotic product was entered in 68 % of the prescriptions and in 98 % of these, the value was *false*. Still, the ability to control narcotic prescriptions and to avoid several simultaneous and overlapping prescriptions to a patient was one of the implementation goals of the e-prescription services.

The e-prescription is based on the HL7 messaging standards, where the original prescription order and all changes concerning the data content are handled as individual messages. The completeness of prescription data varied between message types. In this paper, we discuss five of the 13 message types: prescription orders, pharmacy's preparation dispensing messages, prescription invalidation messages, prescription correction and renewal messages (see Table 1). Of the data elements, ATC-codes, medicinal preparation type codes, purpose of medication use, and dosage instructions for patient are scrutinized. The numbers cover instances of specific documented data elements in comparison to total number of prescriptions (="All data").

**Table 1.** Examples of e-prescription data completeness (abbreviations: o=order, m=message, r=request)

| Message type | Messages (N) | ATC-codes | Preparation type | Purpose | Dosage instructions |
|---|---|---|---|---|---|
| *Prescription o.* | 20616 | 20109 | 20246 | 14824 | 20616 |
| *Dispensing m.* | 35490 | 34694 | 35098 | -- | -- |
| *Invalidation m.* | 423 | 416 | 413 | 273 | 273 |
| *Correction m.* | 321 | 314 | 315 | 257 | 257 |
| *Renewal r.* | 9711 | -- | -- | -- | -- |

ATC-code is currently used to identify the ordered medication, and it covers 97 % of all prescription orders. Analysis of the orders with missing ATC-code entries revealed that typically these orders concerned medicinal preparations that were manually prepared in the pharmacy. Four remaining orders with missing ATC-codes were prescription orders entered by the pharmacist, when the ordering physician had made the actual order in paper or by phone. The medication preparation type code indicates not only the nature of the preparation, but also other preparation details, such as substance strength, package data, dosage form and so on. This data is available in the Pharmaceutical Database. Preparation type covers 98 % of all prescription orders.

In the interim results, data reliability was evaluated in regards to current data specifications and documenting guidelines. The documenting practices clearly varied between different message types. For example, the purpose of medication use was documented in 72 % of the orders. In comparison, the preparation dispensing messages by pharmacies include ATC-codes and preparation type codes, but not the purpose of use or dosage instructions. Similarly, prescription invalidation messages should include all the prescription data and additionally the reason for invalidation, although our analysis shows that only a part of the prescription data is entered. Reasons for invalidation seem to be in line with the specifications. Most typically the prescription was invalidated because of new patient care decisions (88 %), and to a lower degree because of technical mistakes (8 %). In one case, the prescription invalidation was made due to incorrect information given by a patient. Current specifications state that a valid prescription should be invalidated (or marked as fully dispensed) when the medication is ended or the preparation changed. Prescription correction messages should include the prescription data and additionally document the corrections made. Over 97 % of the correction messages include ATC-codes and preparation type codes,

but only 80 % purpose of use and dosage instructions. Most typically a prescription was updated because of being out of date (84 %), and to a lesser degree because of new patient care decisions (9 %). Reasons of correction were missing in 151 correction messages. The highest amount of invalidations and corrections without physician's reasoning documented all originate from same hospital district.

Prescription renewal can be made in a pharmacy or via citizen's own portal providing access to their health care data. Prescription renewal seem to function solely based on prescription document identifiers as even ATC-codes are missing in the data content although ATC-codes are currently used to link continuous medication. Thus, prescription renewal data is not usable for evaluation or research purposes. Only the quantity of documented renewals is assessable.

## 3. Discussion

Based on our study, e-prescription data is complete when stored. However, the reliability of the data varies between different message types. The prescription order and preparation dispensing messages are the most completely documented. To increase re-usability of e-prescription data, the data structure development would benefit from extending beyond the concept of prescription solely as a document for dispensing medication. Instead, prescription data should be aligned with other structural data regarding patient medication orders. Especially substance quantity and dosage information need to be structured further to support automated system functions, such as calculating the total quantity or translating physicians' dose descriptions to patient-friendly phrases or in using automatic dose distribution services by a pharmacy. Similarly, decision support would benefit from increased structuring of prescription data. Currently, e-prescription data and all other medication data are stored in different archives, but can be utilized for decision support integrated in most patient information systems.

Pharmacotherapy is a key part of patient treatment. Centralized storage of medication data and accessibility of that information in any health care organization will particularly improve patient safety. At the moment e-prescription attempts to cover basic medication data "with patient-understandable" language. The demand for double documenting may be one reason why the purpose can be omitted; the purpose of medication use was documented only in 72 % of the orders and not at all in some message types. The physician documents medication purpose using ICD-10 codes, but currently e-prescription allows only free text entry with layman's terms. The difference between various message types can be also explained with user groups. Regarding dispensing data, information intended to guide patients' use of the ordered medicinal preparation are not needed in pharmacy bookkeeping that covers preparation substance strength, quantity, and package data etc. As the rest of the pharmaceutical data and for example diagnosis codes are available also for patients to browse, the suggested layman's terms can be insufficient even for citizens. Uncommon phrases may cause a patient to misunderstand their instructions, especially if no synonym terminology covering professional and layman terms is made available.

An access to the prescription data is already a benefit for the citizens, as they can now inspect up-to-date prescription information. However, due to technical service implementation and user right issues, especially the dosage information is not necessarily up-to-date, which can cause uncertainty to a citizen inspecting the data via

the citizen portal. A typical cause for this is that a physician has updated dosage information into the EHR system used in the organization, but not into the Prescription Centre that is a separate service. Similarly, when nurses update medication data by the order of a physician into a local EHR, they have limited user rights in the Prescription Centre, which causes the e-prescription data to be out of date.

Quality differences between hospital districts are generated by different documenting practices and patient information systems in use. For example, there were variations in documenting permanent medication and preparation substance strength. Messaging standard types increased the differences between data content and lessened the comparability of the data. Our results suggest that reliability of e-prescription data could be increased by further user training. National level guidelines for structured documenting should be emphasized and promoted to decrease differences between hospital districts and professional groups. At the same time, the amount of structured data content should be increased in comparison to free text data entry. With the re-use focus, medication data constitutes a core in decision support and adverse event tools integrated in EHR systems. Additionally, prescription data that covers also dispensing message data provides new potential for evaluating patients' real use of medication as well as commitment to care.

From a patient's point of view, the data content in the citizens' portal should be developed further, besides which the quality of the "translated" data content should be evaluated. Citizens need more diverse tools within the portal; their needs are not met only by making information accessible, but instead tools for entering their own health data would provide more up-to date information also to the physician. However, more explorations are needed regarding citizens' needs and also regarding meaningful use of citizen generated information for the primary and secondary use purposes in health care. Consequently, better interactive and citizen-centred e-health services could be based on such confirmed benefits.

## References

[1] Adler-Milstein, J., Ronchi, E., Cohen, G.R., Pannella Winn L.A. and Jha, A.K. Benchmarking health IT among OECD Countries: Better Data for Better Policy. *Journal of the American Medical Informatics Association* **21**(2014), 111-116.

[2] Vuokko, R., Mäkelä-Bengs, P., Hyppönen, H., Lindqvist, M. and Doupi, P. Impacts of Structuring the Electronic Health Record: Results of a Systematic literature Review from the Perspective of Secondary Use of Patient Data. Accepted for journal publication.

[3] Hyppönen, H., Saranto, K., Vuokko, R., Mäkelä-Bengs, P., Doupi, P., Lindqvist, M. and Mäkelä, M. Impacts of structuring the electronic health record: A systematic review protocol and results of previous reviews. *International Journal of Medical Informatics* **8**(2014), 159-169.

[4] Cheung, N. T., Fung, V., Chow, Y.Y. and Tung, Y. Structured data entry of clinical information for documentation and data collection. *Studies in Health Technology & Informatics* **84**(2001), 609-613.

[5] Hripcsak, G., Knirsch, C., Zhou, L., Wilcox, A. and Melton, G.B. Using discordance to improve classification in narrative clinical databases: An application to community-acquired pneumonia. *Comput Biol Med*. **37**(2007), 296-304.

[6] Hagen, E.M., Rekand, T., Gilhus, N.E. and Gronning, M. Diagnostic coding accuracy for traumatic spinal cord injuries. *Spinal Cord*. **47**(2009), 367-371.

[7] Weiskopf, N.G., Hrispcsak, G., Swaninathan, S., and Weng, C. Defining and measuring completeness of electronic health records for secondary use. *Journal of Biomedical Informatics* **46**(2013), 830-836.

This page intentionally left blank

# Ontology and Decision Support

This page intentionally left blank

# Remote Monitoring of Cardiac Implantable Devices: Ontology Driven Classification of the Alerts

Arnaud ROSIER [a,b,d,1], Philippe MABO [a,c], Lynda TEMAL [e], Pascal VAN HILLE [e], Olivier DAMERON [f], Louise DELEGER [g], Cyril GROUIN [g], Pierre ZWEIGENBAUM [g], Julie JACQUES [h], Emmanuel CHAZARD [i], Laure LAPORTE [j], Christine HENRY [j], and Anita BURGUN [d,k,1]

[a] *INSERM CIC-IT 804, CHU Pontchaillou, Rennes, France*
[b] *Hôpital Privé Jacques Cartier, Groupe GDS, Massy, France*
[c] *Cardiology Department, CHU Pontchaillou, Rennes, France*
[d] *INSERM UMR_S 1138 Eq 22, Paris Descartes University, France*
[e] *LTSI (Inserm UMR1099), Université de Rennes 1, Rennes, France*
[f] *University of Rennes 1, UMR 6074 IRISA, 35042, Rennes, France*
[g] *LIMSI-CNRS, 91405 Orsay, France,* [h] *ALICANTE SA, Lille, France*
[i] *CERIM, Lille University, Lille, France*
[j] *SORIN CRM SAS, Parc Noveos, Clamart, France*
[k] *Hôpital Européen Georges Pompidou, AP-HP, Paris, France*

**Abstract.** The number of patients that benefit from remote monitoring of cardiac implantable electronic devices, such as pacemakers and defibrillators, is growing rapidly. Consequently, the huge number of alerts that are generated and transmitted to the physicians represents a challenge to handle. We have developed a system based on a formal ontology that integrates the alert information and the patient data extracted from the electronic health record in order to better classify the importance of alerts. A pilot study was conducted on atrial fibrillation alerts. We show some examples of alert processing. The results suggest that this approach has the potential to significantly reduce the alert burden in telecardiology. The methods may be extended to other types of connected devices.

**Keywords.** Artificial intelligence, Cardiac implantable electronic devices, Decision support systems, Remote monitoring.

## Introduction

The definition of the Internet of Things (IoT) mainly derives from the vision proposed by Ashton in 1999 [1]: "If we had computers that knew everything there was to know about things - using data they gathered without any help from us - we would be able to track and count everything, and greatly reduce waste, loss and cost. We would know when things needed replacing, repairing or recalling, and whether they were fresh or past their best". Cardiac implantable electronic devices (CIED), including pacemakers,

---

[1] Corresponding Author.

cardiac resynchronization therapy devices and implantable cardioverter defibrillators, are usually not classified at first sight as belonging to the "IoT". Though they were primarily medical devices, CIEDs gained wireless connectivity ten years ago and share with more commonly used connected devices several characteristics that were first described by Ashton: (1) they gather data without any help from us, (2) they are able to track any change in cardiac rhythm and count the episodes of arrhythmia in a patient during a period of time, (3) they reduce costs as they can reduce the number of unnecessary hospital referrals and admissions [2], (4) they help physicians know when replacing of a lead or replacing of the device is required, (5) they improve patient comfort and daily life and have a great impact on the autonomy of the citizen suffering from cardiac disease.

During last decades, new CIEDs were designed from cardiac pacemakers used to treat bradycardia to implantable defibrillators (sudden death prevention) and cardiac resynchronization therapy (standard treatment for many heart failure patients). Efficient systems are required to cope with the higher prevalence of chronic cardiac disorders and heart failure conditions associated with increased life expectancy (e.g., more than 650 000 patients have an implantable defibrillator in the USA [3]).

Remote monitoring of CIED has the potential to make follow-ups more efficient and more effective. It is one of various promising applications of E-Health that will help healthcare systems to meet the increasing needs of its ageing population against the background of budget constraints and capacity shortages.

Nevertheless, telemonitoring is still confronted with many obstacles that prevent full adoption of this innovation – and one of these is the fact that the CIEDs send a lot of remote alerts about arrhythmias to the physicians. The numerous alerts sent by the CIEDs cause a burden for physicians who have to assess their relevance and emergency level: it is estimated that the mean number of event per patient per month is 0.6 [4], raising organisational issues.

The objective of the AKENATON project, funded by the French Research Agency, was to provide a framework for integrating the data transmitted by the CIEDs with their clinical context, to improve alert triage. We designed a prototype for atrial fibrillation (AF) alerts, the more frequent notifications in CIED remote monitoring [4]. This article presents how the ontology supports data integration and reasoning.

## 1. Methods

### 1.1. Scenario

Atrial Fibrillation (AF) is associated with a substantial risk of mortality and morbidity from stroke and thrombo-embolism (TE) that is not homogeneous among patients. Various clinical features have been identified to help stratify risk into high-, intermediate-, or low-risk categories: it has been shown that congestive heart failure, hypertension, age >75, diabetes, prior stroke/transient ischemic attack, and more recently, that female gender, vascular diseases, including myocardial infarction, peripheral artery disease, and complex aortic plaque, all increase thrombo-embolic risk in AF. The CHADS2VASc2 score based on all these parameters is commonly used for risk stratification [5]. Anticoagulation with vitamin K antagonists or new oral anticoagulation drugs (such as rivaroxaban, apixaban, and dabigatran), is prescribed to high- or intermediate-risk patients as it reduces this risk by two-thirds [6]. AF notifications by

manufacturers summons physicians to evaluate device data available on several secure websites and thus represent a high medical burden in terms of alert management.

The overall approach consists in classifying the alerts transmitted by the CIEDs so that the situations that require a medical advice be submitted to the cardiologist according to their respective rankings: critical, urgent, moderate, not urgent. In order to classify the alerts according to their clinical importance – including TE risk –, the system takes into account the patient's context and extracts relevant data from the patient's electronic medical record (EMR).

### 1.2. Domain ontology

Several terms and concepts could be extracted from existing terminologies such as the NCI Thesaurus, Human Disease Ontology (DOID), SNOMED-CT. However, the coverage was extremely low (less than 40%) for the following domains: cardiac pacing, electrophysiology, cardiac devices. Therefore, we designed the Cardio-Vascular Disease Ontology (CVDO) to serve as a basis for providing formal definitions of the domain entities. The ontology has been made available as part of the OBO Foundry, a repository for formal open source ontologies [7]. The ontology was extended to provide specific entities related to CIEDs [8]. Formal definitions were provided for many concepts of the AF scenario so that the application ontology could support reasoning.

For each alert, the ontology is instantiated with the patient data: (1) the data extracted from the patient EMR is integrated; (2) the patient's CHADS2VASc score is determined from the data that have been integrated; (3) the alert is classified. We give examples of such ontology-based reasoning in the Results section. A prototype was developed and evaluated on 1783 AF alerts from a retrospective cohort of sixty patients from the PARADYM cohort over a period of one year [9]. Ontology-driven alert classification was blindly assessed by two physicians, including a cardiologist.

## 2. Results

The application ontology is an OWL2 ontology (Web Ontology Language version 2) [10]. This ontology comprises 252 classes and 25 relations, among which 17 ObjectProperties and 8 DataProperties. It also contains a total of 1112 axioms that are needed for reasoning tasks. Forty classes are defined classes with necessary and sufficient conditions used for classifying the alert cases. 196 disjunction axioms were implemented.

Figure 1 illustrates patient instantiation including a list of devices, diseases, events, medications and CHADSVASc score items (on the left), and two types of closure that were necessary for reasoning, namely the assumptions that support the closure to address the open world assumption in OWL, and the unique name assumption (i.e., instances are distinct), which is missing by default in OWL.

In Figure 2 and 3 are displayed examples of how formal logical definitions support inferring CHADSVASc scores. All the reasoning steps are made available (Figure 3). Only one Semantic Web Rule Language (SWRL) rule was implemented, namely to sum all the elementary scores to obtain the CHADSVASc global score [11]. 1749 out of 1783 alerts (98%) were adequately classified by the ontology-based method. No error was due to ontology reasoning per se: all errors were related to concept extraction from EMRs. For instance, a case of new diagnosis of diabetes mellitus and hypertension could not be

captured but were considered by physicians based on the HBA1C and arterial tension values, despite they didn't meet strict diagnosis criteria Also, false positives of stroke were captured in medical justification for radiographic exams.
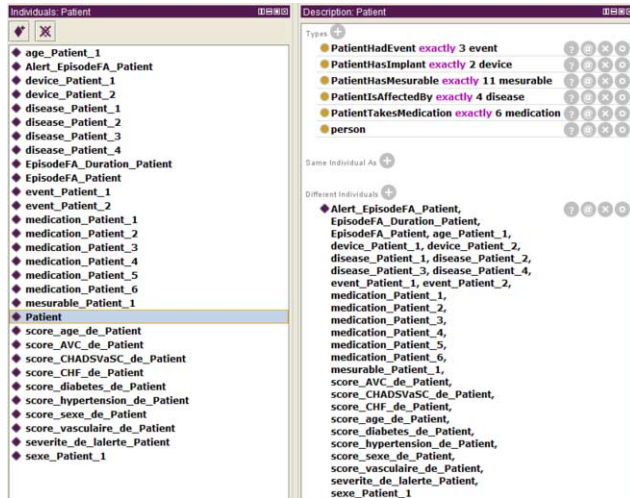


**Figure 1.** Patient instantiation.



**Figure 2.** Critical alert formal definition (upper right) based on patient anticoagulation status, AF episode duration and CHADSVASc score.
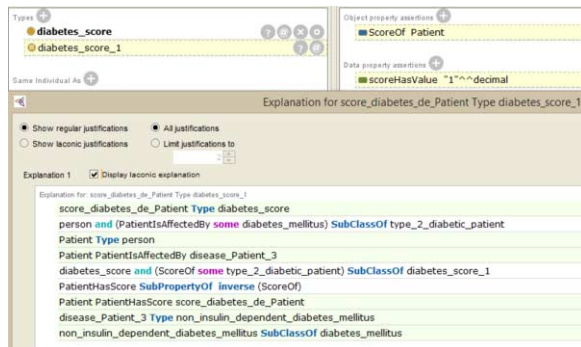


**Figure 3.** Inference of diabetes score for a patient with type 2 diabetes, with inference steps.

84% of the alerts were classified as "not urgent". It suggests that our approach possibly dramatically decrease the alert burden for the physicians. The classification process took an average 450ms per patient using the OWL reasoner (Pellet).

## 3. Discussion

Theoretically, the use of CIEDs can ensure better follow up of potentially life-threatening situations such as arrhythmias. In practice, the growing number of recipients of CIEDs and other connected devices will lead to rapidly increasing workload for the follow-up of these patients. With the objectives of remotely monitoring a patient, the data transmitted by the CIED must be regarded as part of the patient record, and must be combined with the EMR data to guide medical decision. We have shown that the classification of AF alerts was feasible and accurate. This demonstrates that OWL- based ontology modelling techniques can reliably perform the reasoning necessary to propose a severity level associated with CIED alerts. Moreover, the performances are compatible with the scalability requirements of telecardiology that motivated the AKENATON project.

This example illustrates a novel approach that may be extrapolated to other connected devices and is likely to become much more needed with the development of "the IoT" and remote monitoring, and with the widespread adoption of EMRs.

A challenging issue is to develop methods to extract information from EMRs, as well as formal ontologies and ontology driven algorithms in the domains that require information triage. Such efforts are needed to assist medical experts in alert triage and clinical decision tasks.

## References

[1]    Ashton, K. (2009). That "Internet of Things" Thing - RFID Journal. RFID Journal. Retrieved from http://www.rfidjournal.com/articles/view?4986

[2]    Daubert JC. et al. 2012 EHRA/HRS expert consensus statement on cardiac resynchronization therapy in heart failure: implant and follow-up recommendations and management. *Europace*, 2012, **14**(9), 1236–86.

[3]    Brown D.W. et al. Trends in Hospitalization for the Implantation of Cardioverter-Defibrillators in the United States, 1990–2005. *Am J Cardiol.* 2008 Jun 15;**101**(12):1753-5.

[4]    Lazarus A. Remote, wireless, ambulatory monitoring of implantable pacemakers, cardioverter defibrillators, and cardiac resynchronization therapy systems: analysis of a worldwide database *Pacing Clin Electrophysiol*. 2007 Jan;30 Suppl 1:S2-S12.

[5]    Lip G.Y.H., et al. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest*, 2010, **137**(2) : 263–72.

[6]    Aronis KN, et al. Paroxysmal atrial fibrillation and the hazards of under-treatment. *Int J Cardiol*. 2015 Sep 11;202:214-220.

[7]    Barton, A., Rosier, A. Burgun, A & Ethier, J.-F. The Cardiovascular Disease Ontology, *Proceedings of the 8th International Conference on Formal Ontology in Information Systems (FOIS),* 409-414, P. Garbacz & O. Kutz (Eds.), Amsterdam: IOS Press.

[8]    Rosier A, Mabo P, Chauvin M, Burgun A. An ontology-based annotation of cardiac implantable electronic devices to detect therapy changes in a national registry. *IEEE J Biomed Health Inform*. 2015 May;**19**(3):971-8

[9]    Paradym cohort [http://clinicaltrials.gov/ct2/show/NCT01169246].

[10]   W3C OWL Working Group. (2012). OWL 2 Web Ontology Language Document Overview. OWL 2 Web Ontology Language, 1–7. Retrieved from http://www.w3.org/TR/owl2-overview/

[11]   Horrocks, I., Patel-schneider, P. F., Boley, H., Tabet, S., Grosof, B., & Dean, M. (2004). SWRL : A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission 21.

# Ontology-Oriented Programming for Biomedical Informatics

Jean-Baptiste LAMY[a,1]

[a] *LIMICS, Université Paris 13, Sorbonne Paris Cité, 93017 Bobigny, France,*
*INSERM UMRS 1142, UPMC Université Paris 6, Sorbonne Universités, Paris*

**Abstract.** Ontologies are now widely used in the biomedical domain. However, it is difficult to manipulate ontologies in a computer program and, consequently, it is not easy to integrate ontologies with databases or websites. Two main approaches have been proposed for accessing ontologies in a computer program: traditional API (Application Programming Interface) and ontology-oriented programming, either static or dynamic. In this paper, we will review these approaches and discuss their appropriateness for biomedical ontologies. We will also present an experience feedback about the integration of an ontology in a computer software during the VIIIP research project. Finally, we will present OwlReady, the solution we developed.

**Keywords.** Knowledge representation (computer), Computer programs and programming, Ontology, Ontology-oriented programming, Biomedical knowledge, Open-world assumption

## Introduction

The biomedical domain is one of the most complex domain of Human knowledge today. It is necessary to structure and formalize this knowledge adequately. The field of knowledge representation leads to the development of ontologies, which can link knowledge together and produce inference using a reasoner. Ontologies can represent universal statements (true for all individuals of a given class, e.g. medical knowledge about disorders or drugs), terminological statements (related to terms in a given natural language, e.g. synonyms for disorder names) and assertional statements (related to a given individual, e.g. medical data of a given patient) [1].

Many methods and tools have been proposed for the design, maintenance, alignment or evaluation of biomedical ontologies [2]. However, fewer options are available for ontology programming interface, another problem identified by A. Rector et al. [3]: how to access and manipulate an ontology in a computer program, for example to connect the ontology to a database or to generate a website from the inferences produced by the ontology?

In this paper, we will first describe some particularities of ontologies in the biomedical domain. Then, we will review the various approaches proposed for accessing ontologies in a computer program, and we will discuss which approach is the most appropriated for biomedical ontologies. Finally, we will give an experience

---

[1] Corresponding Author.

feedback about the integration of an ontology in computer software during a research project, and we will present OwlReady, the solution we developed.

## 1. Methods

Ontology in biomedical informatics. An ontology is the specification of the concepts, their attributes and relationships, in a given domain of discourse, for instance using the Ontology Web Language (OWL). Ontologies can be used for performing logical inferences and linking knowledge together in the semantic web. Ontologies rely on the open-world assumption, i.e. any fact is considered as possible until it has been explicitly stated that the fact is impossible.

The inherent complexity of the biomedical domain leads to some particularities when designing ontologies. (a) The open-world assumption is not always appropriate for medical reasoning. It is desirable for patient-related knowledge, for example, it allows reasoners to make hypotheses about unknown patient's disorders, e.g. in a diagnostic decision support system. On the contrary, it is not appropriate for drug- or disorder-related knowledge, for example, when considering drug adverse effects, we usually consider that all adverse effects are known (even if this may not be completely true) and that any adverse effect that is not known cannot occur. Thus a decision support system is expected to reason only on known effects and not to make hypotheses about unknown potential adverse effects. This would lead to stupid alert message like "the patient has hyperkalaemia; the drug you are prescribing does not cause hyperkalaemia but it might have an unknown adverse effect of hyperkalaemia".

(b) Many medical concepts cannot be represented by individuals but need classes to represent them. Disorders can be expressed at various levels of granularity (e.g. inflammatory bowel disorders, Crohn disease, severe Crohn disease with skin manifestation...) with inheritance (is a) relations between them. The same problem occurs for Human-created artefact such as drug treatments or medical acts. Drugs can be described by chemical or therapeutic classes, by active principles, by brand names, or even with a dose or an indication (e.g. aspirin in the antiplatelet indication). Consequently, both disorders and treatments should be classes rather than individuals in medical ontologies.

### 1.1. Traditional API for OWL vs ontology-oriented programming.

Two approaches have been proposed for accessing an ontology in a computer program. (a) Traditional API (Application Programming Interface) for OWL, such as OWL API [5] in Java, provide functions and classes for manipulating OWL constructs, e.g. an OWL class is an instance of OWLClass from the programmer's point of view. (b) Ontology-oriented programming tries to unify the ontology with the object model of the programming language, e.g. an OWL class is a class from the programmer's point of view. This approach exploits the similarities between ontologies and the object-oriented programming paradigm [6]: classes, properties and individuals in ontologies correspond to classes, attributes and instances in object models.

Ontology-oriented programming leads to shorter and more readable source codes, as shown by W3C [4] (Figure 1), and it is well-known in software development that a shorter code means fewer bugs. This is particularly interesting in biomedical

informatics, because we have seen that treatments and disorders are classes, and classes are even more complex to manipulate than instances.

**Traditional API with Java:**

```
public static float getOrderCost(OWLIndividual order) {
    OWLModel model = order.getOWLModel();
    OWLProperty drugProperty  = model.getOWLObjectProperty("drug");
    OWLProperty priceProperty = model.getOWLDadatypeProperty("price");
    float cost = 0.0;
    Iterator drugs = order.listPropertyValues(drugProperty);
    while(purchases.hasNext()) {
        OWLIndividual drug = (OWLIndividual) drugs.next();
        Float price = (Float) drug.getPropertyValue(priceProperty);
        cost = cost + price.floatValue();
    }
    return cost;
}
```

**Ontology-oriented programming with Python and OwlReady:**

```
class Order(Thing):
    def get_cost(order):
        cost = 0.0
        for drug in order.drugs: cost = cost + drug.price
        return cost
```

**Figure 1**. Two examples of ontology integration in a computer program. Both examples compute the total cost of a drug order (considering one box of each drug).

## 1.2. Static vs dynamic ontology-oriented programming

Two approaches exist for ontology-oriented programming. (a) The static approach consists of software that generates the source code for classes, from an ontology described in OWL. Modules have been implemented for Java [7] and C# [8]. They allow access to the ontology and to verify typing at compile time, but due to their static nature, they do not allow inference, classification or dynamic class creation. More recently, a semi-dynamic approach in Java [9] allows inference on individuals but not on classes. (b) The dynamic approach uses dynamic programming languages to generates classes and instances from the ontology at run time. In this approach, the same class is considered from an ontological point of view (following the open-world assumption) and from an object-model point of view (following the closed-world assumption, i.e. any fact is considered as impossible until it has been explicitly stated that the fact is true). A first module was proposed in Common Lisp [6] and a limited prototype in Python [10].

We have seen previously that, in the biomedical domain, both open-world and closed-world assumption are desirable, and that manipulating and classifying classes is a requirement. Consequently, the dynamic approach seems the right one for biomedical informatics.

## 2. Results

Experience feedback and contribution. The VIIIP (Integrated Visualization of Information about Pharmaceutical Innovation) research project aims at presenting information about new drugs to physicians. To guarantee an independent information,

this information is produced automatically by comparing the new drug to the older ones, using criteria such as efficacy in clinical trials, contraindications and known adverse effects.
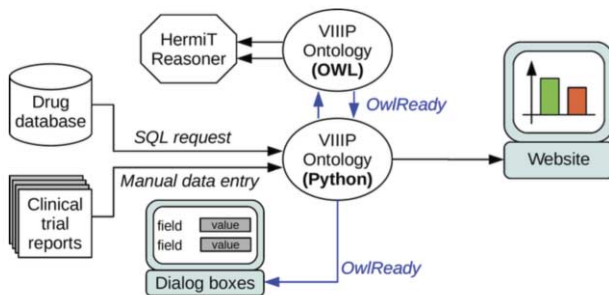


**Figure 2**. General structure of the VIIIP platform.

The automatic comparison of drug properties like contraindications is not easy because contraindications are often expressed at different levels of granularity in drug databases, e.g. rhythm disorder vs risk of torsades de pointes. To perform the comparison, we designed an ontology of contraindications. The ontology is populated from the French drug database Thériaque (http://theriaque.org), then a reasoner computes inferences, and finally the resulting inferences are presented in a website.

The general structure of the project was clear, but we experienced difficulties for connecting the ontology to the database and the website, and more generally to manipulate the ontology in the computer program. These difficulties lead us to a reflection about methods for accessing ontologies, and to the development of OwlReady (https://pypi.python.org/pypi/Owlready), a Python 3 module for ontology-oriented programming with full class-support, including dynamic class creation and classification of classes at run time using the HermiT reasoner [11]. OwlReady supports OWL 2.0. An experimental feature allows to automatically generate dialog boxes for editing individuals and classes in the ontology.

We successfully used OwlReady for implementing the VIIIP platform (Figure 2). Populating the ontology from the results of SQL requests, calling the reasoner, and generating HTML pages from the ontology were easy, from a computer-science point of view (but we encountered other problems related to the quality of data, out of the scope of this paper). The automatic generation of dialog boxes was very convenient, it allowed to modify the ontology without having to manually update the editing interface. The computation time of ontology-oriented programming with OwlReady and Python was higher than a traditional API in Java. However, the difference is insignificant compared to the time consumed by the HermiT reasoner, or the time we gained during software development (shorter source code implies faster development).

## 3. Discussion

We have shown that, in the biomedical domain, dynamic ontology-oriented programming is an interesting approach for integrating ontologies in computer software. In facts, it leads to simpler and shorter source code, while computer programs manipulating biomedical ontologies tend to be complex since disorders and treatments are represented by classes and not instances. It also allows a "mix" of open- and closed-

world assumption. We tested this approach in a research project and we proposed OwlReady, a module for dynamic ontology-oriented programming in Python.

In its current development stage, OwlReady is really practical for using ontology for a reasoning purpose, typically when one needs to dynamically create concepts, perform some reasoning on them, and then present the result of the reasoning. However, due to the lack of support of ontology file-formats (currently limited to a fair subset of OWL / XML), it is not yet well-suited for linking knowledge together.

Future works on OwlReady will focus on the use of classes. For example, some class restrictions could be exposed as if they were properties of the class, in a similar way to individual properties (e.g. the "Class property_x value 1" OWL restriction would be equivalent to "Class.property_x = 1" in Python). Class-class relations (i.e. all individuals of a class are related to all individuals of another, for example "all drugs of the anti-vitamin K pharmacological class interact with all drugs of the NSAI (Non-Steroid Anti-Inflammatory) pharmacological class") are often problematic in OWL because they cannot be represented directly. This is another area of possible improvement.

## Acknowledgment

## References

[1] Schulz S, Jansen L. Formal ontologies in biomedical knowledge representation. Yearb Med Inform. 2013;8:132–46.

[2] Yu AC. Methods in biomedical ontology. J Biomed Inform. 2006;39(3):252–266.

[3] Rector A, Horridge M, Iannone L, Drummond N. Use Cases for Building OWL Ontologies as Modules: Localizing, Ontology and Programming Interfaces & Extensions. In: 4th Int Workshop on Semantic Web enabled software engineering (SWESE-08); 2008.

[4] Knublauch H, Oberle D, Tetlow P, Wallace E. A Semantic Web Primer for Object-Oriented Software Developers. W3C Working Group Note. 2006.

[5] Horridge M, Bechhofer S. The OWL API: A Java API for OWL ontologies. Semantic Web 2. 2011;p. 11–21.

[6] Koide S, Aasman J, Haflich S. OWL vs. Object Oriented Programming. In: the 4th International Semantic Web Conference (ISWC 2005), Workshop on Semantic Web Enabled Software Engineering (SWESE); 2005. p. 1–15.

[7] Kalyanpur A, Pastor D, Battle S, Padget J. Automatic mapping of OWL ontologies into Java. In: Proceedings of the Sixteenth International Conference on Software Engineering & Knowledge Engineering (SEKE'2004); 2004. p. 98–103.

[8] Goldman NM. Ontology-oriented programming: static typing for the inconsistent programmer. In: Lecture notes in computer science: the SemanticWeb, ISWC. vol. 2870; 2003. p. 850–865.

[9] Stevenson G, Dobson S. Sapphire: Generating Java Runtime Artefacts from OWL Ontologies. In: Lecture Notes in Business Information Processing, Advanced Information Systems Engineering Workshops. vol. 83; 2011. p. 425–436.

[10] Babik M, Hluchy L. Deep Integration of Python with Web Ontology Language. In: Proceedings of the 2nd workshop on scripting for the semantic web. Budva, Montenegro; 2006. p. 1–5.

[11] Motik B, Shearer R, Horrocks I. Hypertableau reasoning for description logics. Journal of Artificial Intelligence Research. 2009;36:165–228.

# An Integrated Children Disease Prediction Tool within a Special Social Network

Marika APOSTOLOVA TRPKOVSKA[a,1], Sule YILDIRIM YAYILGAN [b],
and Adrian BESIMI[a]

[a] *South East European University, Ilindenska n. 335, 1200 Tetovo, Macedonia*
[b] *Norwegian University of Science and Technology, 2815 Gjøvik, Norway*

**Abstract.** This paper proposes a social network with an integrated children disease prediction system developed by the use of the specially designed Children General Disease Ontology (CGDO). This ontology consists of children diseases and their relationship with symptoms and Semantic Web Rule Language (SWRL rules) that are specially designed for predicting diseases. The prediction process starts by filling data about the appeared signs and symptoms by the user which are after that mapped with the CGDO ontology. Once the data are mapped, the prediction results are presented. The phase of prediction executes the rules which extract the predicted disease details based on the SWRL rule specified. The motivation behind the development of this system is to spread knowledge about the children diseases and their symptoms in a very simple way using the specialized social networking website www.emama.mk.

**Keywords.** Semantics, Social Networking, Ontology, Disease Prediction

## Introduction

Today more often new diseases are identified and diagnosed. On daily basis children become patients and are diagnosed with various diseases. A disease is identified by examining its symptoms and signs which human body develops. The cause for one disease could be different from the cause of another one. However, a general cause for most of the diseases is unawareness. This is caused by the lack of knowledge about symptoms and signs developed by human body and the lack of knowledge about diseases. To tackle this problem, firstly there is a need for a general knowledge tool that can help us with sharing knowledge about children diseases and their symptoms and signs. Ontology is a new paradigm and technology that is used primarily for sharing knowledge in digital form that is understandable by computers. Secondly, using this technology, we developed a disease prediction model for children. We identified all the symptoms and signs that correspond to a disease (which symptoms must appear and which are optional) by extracting these relationships from the official online medical data and consulting with a paediatrician before they were inserted in the developed ontology.

Although there exist many medical systems that somehow benefit from ontology use, yet very few of them deal with children diseases [12]. We have selected children diseases as our focus area firstly to target the parents and secondly the children.

---

1 Corresponding Author.

Healthy child usually means a healthy adult. Every stage of human life is important, but the biggest health related mistakes are made in childhood [14]. Besides, health complications which occur during childhood are the cause and the predisposition for disease in later life [13]. A lot of evidence promoting the use of information and communication technologies in health care exists [5]. Ontology and a rule based intelligent information system to detect and predict myocardial diseases was implemented by Antonio and presented in his paper [2]. The system was used in pre-hospital health emergencies, remote monitoring of patients with chronic conditions and medical collaboration through sharing health-related information. Another author, Maja, proposed an ontology based information system for biomedical field [3]. In her next paper [4], an ontology based multi agent system to support human disease study and control is introduced. A coherent biomedical literature clustering and summarization approach that takes advantage of ontology-enriched graphical representations significantly improved the quality of document clusters and understandability of documents through summaries as introduced by Illhoi [6]. In 2006, Hongyi brought in a method for obtaining biological functions of genes by using the GENE ontology method [7]. Akifumi evaluated existing medical ontologies and proposed future directions for a medical knowledge repository system with three knowledge repositories [8]. Jaszuk constructed a reasoning mechanism that diagnoses the possible diseases by comparing the data obtained for a patient with the semantic models of diseases [9]. A specialized ontology in biomedical information storage and processing had been carried out by Tharam and presented in his paper [10].

Our research proposes a specially designed social network (www.emama.mk) with an integrated specialized tool for predicting children diseases. Contrarily to the existing work, we provide a simple and clear prediction methods that concentrate on a several variables (not just symptoms or signs). Mothers may stay in touch, share parenting tips, create groups to discuss various topics, activate various health related services like a vaccination reminder by using the proposed social network based prediction tool. It will especially help new mothers living in rural areas, not only to exchange experiences with other mothers, but also to identify what disease their children might have. We focus on diseases that are most common and specific in Macedonia in the early childhood which are respiratory infections, diarrhea, infections of the urinary tract and skin, and a wide spectrum of allergies. Only registered users can activate the prediction system [1]. We emphasize that the disease prediction tool is only for providing initial supportive information and additional consultation with medical specialist is mandatory in order to determine the actual disease and to continue with the appropriate treatment [11].

## 1. Methods

Our model for developing the prediction tool is aimed at optimizing the treatment of large amounts of data relevant for the diagnosis search (searching for the disease which is most relevant to the given symptoms in the ontology), while striving to obtain the probability of possible results with a very high accuracy. Having multiple data sources and multiple search techniques, it provides a broader view of the search problem (knowing as much information as possible about the patient related to its illness), and produces more relevant results. A paediatrician is consulted to identify the type of information required to predict a disease. A questionnaire is provided in the proposed

social network site which consists of five questions (related to child age, gender, body temperature, symptoms and vaccinations) formed in collaboration with the paediatrician. This questionnaire is in fact a profile for the child that we need to store for further processing.

At the first stage, before applying the model developed, it is necessary that the system gathers the required information and stores it in a MySQL database. The CGDO ontology plays a key role in developing the prediction tool. It uses patients' data taken from the questionnaires in combination with the earlier acquired and stored data related to children diseases, including disease classes, object or data properties, data type definitions, assertions, annotations, axioms and SWRL rules in required formats. The second stage of ontology development is composed of two steps. In step one, the class hierarchy, the objects and data properties are defined. In step two, rules are derived from the disease details and defined axioms. A prediction rule is given in Fig. 1.

```
Person(?p) ∧ has_Case(?p, Patient_Case_101) ∧ has_Prevention(?p, ?pr) ∧ Disease(?d) ∧
is_Prevention_for(?pr, ?d) ° sqwrl:makeSet(?s1, ?d) ° sqwrl:element(?e, ?s1) ∧
has_Diagnose(Patient_Case_101, ?d2) ∧ sqwrl:makeSet(?s2, ?d2) ∧ sqwrl:difference(?s3, ?s2, ?s1) ∧
sqwrl:element(?e2, ?s3) ∧ Disease(?e2) ∧ has_Symptom(?e2, ?s) ∧ has_Symptom(Patient_Case_101, ?s)
→ sqwrl:select(?p, ?e2) ∧ sqwrl:count(?s) ∧ sqwrl:orderByDescending(?s) ∧
sqwrl:columnNames("Patient_case","Disease", "Factor")
```

**Figure 1.** Rule for disease prediction by eliminating the diseases for which the patient is vaccinated.

The next stage requires a mapping from the database to the developed CGDO ontology. We focus on database-to-ontology mapping approach which considers that ontology and a legacy database already exist. The goal is to create a mapping and/or populate the ontology by the database contents. Mapping is more complex because different levels of overlap between the database domain and the ontology domain may be found. Those domains do not always coincide because the modelling criteria used for designing databases are different from those used for designing ontology models. Thus, in the further stage of this model, a conversion of the database into RDF notation is done by the help of D2R server and by applying D2RQ mapping language. An OWL document, containing data from the database, in order to be imported into the existing CGDO ontology, is created by using the Protégé tool. The prediction stage of the model development is grounded of a semantic reasoner and SWRL rules. The semantic reasoner is used for checking the consistency of the relationship among the classes and their properties. This is done in order to validate the SWRL rules. One of the existing SWRL rules is created from valid relationships to predict a disease from the given symptoms (Fig. 1). Here with a help of SQWRL (SWRL-based language for querying OWL ontologies) we created SQL-like operations to retrieve knowledge inferred by the SWRL rules. SQWRL helped us to effectively build a query language on top of the existing SWRLs. The results for the predicted diseases are stored as instances in the ontology for later use [13].
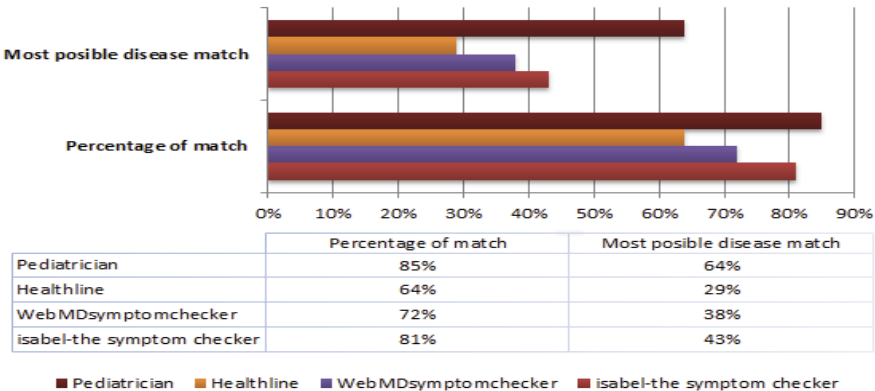
## 2. Model testing and Results

In order to test the proposed prediction, and especially the rules for predicting possible diseases (Fig. 1), the proposed method was applied on 100 patient cases (questionnaire data available from 100 test cases) and the results were compared with three other electronic tools available for predicting diseases: *isabel-the symptom checker*,

*WebMDsymptomchecker* and *Healthline* as well as with predictions of the paediatrician. None of the three tools are using ontologies.

An example case concerns the prediction of acute_diarrhea by applying our method. Based on the available patient data from the questionnaire; appeared symptoms: *Abdominal cramp, myalgia, nausea and watery stool,* age: *Infant*, gender: *Male* and region: *Western Europe*, the system predicted **acute_diarrhea** as the most probable disease. By applying the rule in Fig. 1, from all the diseases in the CGDO ontology, only four possible diseases are selected, namely: acute_diarrhea, diarrhea, chickenpox and campylobacteriosis with the possibility factors of 64, 48, 32, and 16 respectively. The factor is calculated by using the number of matched symptoms, in this case respectively 4, 3, 2 and 1 symptoms for the four diseases are related to the symptoms available for the predicted disease in the CGDO ontology. Acute_diarrhea appears as a possible common disease and in the results of Isabel Symptom Checker, and in the results from the WebMD symptom checker, it is not appearing at all, but in Healthline and in paediatrician predictions, it is diagnosed as a possible disease.

In Table 1, we see the *percentage of match* and the *most possible disease match* regarding our method and the other four prediction sources. On one hand, the percentage of match gives the percentage of diseases similarly predicted by our method and the other prediction tools. On the other hand, the most possible disease match indicates the percentage of concordance of the disease that has the highest probability with the actual disease of the patient. We conclude that our predictions are the closest to the ones of paediatricians due to the fact that all the data available in the ontology are very carefully selected and consulted with a medical specialist for children disease. The other three online tools are big repositories of diseases in comparison to our repository of diseases, however they are not taking paediatricians opinions into consideration.

**Table 1.** Percentage of matching the possible predicted diseases and of the most probable disease



| | Percentage of match | Most posible disease match |
|---|---|---|
| Pediatrician | 85% | 64% |
| Healthline | 64% | 29% |
| WebMDsymptomchecker | 72% | 38% |
| isabel-the symptom checker | 81% | 43% |

## 3. Conclusion

This paper proposes a model for developing an integrated disease prediction tool for children which uses available patients' data and the CGDO ontology. Patients' data are obtained through online questionnaires or profiling.

In order to be fully functional and to benefit from the social network, the system has been developed as part of the specially designed social network called www.emama.mk. The developed prediction system is useful for mothers to become aware of possible children diseases and take rapidly valuable actions when necessary. In addition, this social network system is used as a data collection point for future analysis. The prediction model developed has been tested and shows promising results. The possible limitations of the proposed prediction model are that it produces relatively few outcomes in the form of disease predictions due to a limited number of diseases in the ontology repository, It requires also being evaluated by more external validators e.g. more paediatricians from Macedonia or other countries. Despite the few outcomes, our model performs quite well. These limitations highlight the need for further research in this important but neglected topic. The future work requires "feeding" the ontology with more relevant data and symptoms as well as more user profiles in the social media page. It also includes making further analysis with the available data and conducting qualitative interviews with the paediatricians and mothers about the perceived usefulness of the model.

## References

[1] M.A. Trpkovska, B. Cico, I. Chorbev, *Application of social media in e-health,* In Embedded Computing (MECO), 2014 3rd Mediterranean Conference, IEEE, Budva, Montenegro, (2014), 314 – 317.

[2] A.J. Jara, F.J. Blaya, M.A. Zamora, A.F.G. Skarmeta, *An Ontology and Rule Based Intelligent Information System to Detect and Predict Myocardial Diseases,* In Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine (2009), 5-7.

[3] H. Maja, C. Elizabeth, W. Pronpit, M. Robert, *Disease Ontology based Grid Middleware for Human Disease Research Study,* Proceedings of 30th Annual Conference of the IEEE Industrial Electronics Society, vol no. 1, (2004), 480- 486.

[4] H. Maja, C. Elizabeth, *Ontology-based Multi-agent System support Human Disease Study and Control*, Proceedings of the conference on Self Organization and Autonomic Informatics (2005), 129-141.

[5] A. Díaz-Chao, J. Torrent, D La Casta., F. Saigí, *Improving integrated care: Modelling the performance of an online community of practice,* International Journal of Integrated Care, 14, e007, 2014.

[6] Y. Illhoi, H. Xiaohua, *Biomedical Ontology MeSH improves Document Clustering Qualify on MEDLINE Articles: A comparison study,* Proceedings of the 19th IEEE Symposium on computer –based medical systems (2006), 577-582.

[7] Z. Hongyi, P. Jiexin, Z. Junying, *Construction of gene regulatory networks based on gene ontology and multivariable regression.* Proceedings of International conference on mechatronics and automation (2006), 1324-1328.

[8] T. Akifumi, M. Naoko, M. Hajime, *Medical Ontologies as a Knowledge Repository,* Proceedings of IEEE/ICME international conference on complex medical engineering (2007), 487-490.

[9] M. Jaszuk,, G. Szostek, A. Walczak, *An ontology building system for structuring medical diagnostic knowledge,*Proceeding of the 3rd Conference on Human System Interactions (HIS), Rzeszow, Poland, (2010), 203-210.

[10] S.D. Tharam, C. Elizabeth, H. Maja, *Ontology Support for Biomedical Information Resources*, Proceedings of 21st IEEE International Symposium on Computer-Based Medical Systems (2008), 7-16.

[11] M.A. Trpkovska, B. Cico, *Prototype System for Multiple Sources Multiple Search Techniques Prediction (Algorithm and architecture),* In Proceedings of 2nd International Symposium on Computing In Informatics And Mathematics (ISCIM 13), Epoka University, Tirana, (2013), 84-89.

[12] A. Dahu, Z. Yinsheng, S. Huilin, *Develop of diseases pest information system based on WEBGI*, Proceedings of International Seminar on Future Biomedical Information Engineering, (2013), 461-464.

[13] M.A. Trpkovska, B. Cico, I. Chorbev, *Using Social Media to Predict Children Disease Occurrence*, In Proceedings of 7th IADIS International Conference Information Systems, Spain, (2014), 209 - 216.

[14] NIH MedlinePlus, *Sickle Cell Disease: Symptoms, Diagnosis, Treatment and Recent Developments*, (https://www.nlm.nih.gov/medlineplus/magazine/issues/winter11/articles/winter11pg18.html).

# Ontological Foundations for Tracking Data Quality through the Internet of Things

Werner CEUSTERS[a,1] and Jonathan BONA[a]

[a] *Department of Biomedical Informatics, University at Buffalo, Buffalo, NY, USA*

**Abstract.** Amongst the positive outcomes expected from the Internet of Things for Health are longitudinal patient records that are more complete and less erroneous by complementing manual data entry with automatic data feeds from sensors. Unfortunately, devices are fallible too. Quality control procedures such as inspection, testing and maintenance can prevent devices from producing errors. The additional approach envisioned here is to establish constant data quality monitoring through analytics procedures on patient data that exploit not only the ontological principles ascribed to patients and their bodily features, but also to observation and measurement processes in which devices and patients participate, including the, perhaps erroneous, representations that are generated. Using existing realism-based ontologies, we propose a set of categories that analytics procedures should be able to reason with and highlight the importance of unique identification of not only patients, caregivers and devices, but of everything involved in those measurements. This approach supports the thesis that the majority of what tends to be viewed as 'metadata' are actually data about first-order entities.

**Keywords.** Biological Ontologies, Internet, Metaphysics

## Introduction

Although success stories for the use of electronic health records (EHR) to support individual patient care and biomedical research do exist , others argue that '*EHRs have yet to truly fulfil their promise to support clinicians in their patient care activities, including the essential work of building the patient's story*' . Also that the secondary use of EHR data to support, for instance, comparative effectiveness research is at least cumbersome because '*electronic health record data from clinical settings may be inaccurate, incomplete, transformed in ways that undermine their meaning, of unknown provenance, of insufficient granularity*' [4]. The major bottle neck for appropriate EHR use being the quality of data entry, this involves not only errors in human data entry but also the failure to enter data which are required [5].

It is precisely here that the Internet of Things (IoT) might bring tremendous advantages by avoiding the burden of structured data entry by humans through the connection of devices that use network services to enter data automatically. These devices, some of which not being designed specifically for healthcare purposes, will range from context-aware thermometers, weighing scales and tonometers to intelligent video and sound recorders that during a patient encounter – and if the patient so desires also during his everyday activities and insofar relevant to his health – record every

---

[1]Corresponding Author.

single event or state. Powerful analytics software will then have the capacity to extract all and only meaningful data from these recordings.

But as the use of EHR systems might itself constitute a risk for patient safety, so may the IoT lead to adverse events due to device malfunctioning or communication errors leading to erroneous data entry. However, whereas generally the odds for system malfunction increase relative to the number of devices that are part of the system, the IoT can be set up in such a way that these odds decrease by exploiting the fact that devices can observe and measure not only what is the case for the patient, but also for the patient's environment including the interconnected devices themselves! This requires an IoT for health not only to manage data about the patient but also about its own components and how these components contribute to assertions about the patient.

In this paper, we propose Ontological Realism as a methodology to identify and describe (1) which components within the ontological structure exhibited by the configurations of entities observed and measured by IoT devices are essential and (2) the abstract syntax towards which the output of IoT devices (or the subsequent interpretation thereof) should be formatted, for such devices and their operation to minimize both the burden of data entry and the risks for assertion errors.

## 1. Methods

Ontological Realism (OR) is a theory that defines the principles for high quality ontology development used in the Basic Formal Ontology (BFO) and the ontologies accepted in the Open Biomedical Ontology (OBO) Foundry [6]. Crucial for the proposal advanced here is that OR recognizes two major types of components out of which reality is built: (1) *particulars* such as this paper and its authors – all entities that carry identity, and (2) *universals*, for example those generic entities denoted by general terms such as 'paper', and 'person', which have particulars as instances. Particulars may enjoy *relations* with other particulars so as to form *configurations*. An example is the configuration which constitutes the ground truth for the assertion that this particular paper was the output of a particular collaborative writing process in which the particulars Werner Ceusters and Jonathan Bona, as well as their beliefs about the adequacy of the proposal advanced, all participated during certain time periods.

Referent Tracking (RT) is an OR-based paradigm for knowledge management that originally has been introduced in the context of EHR keeping [7]. Whereas realism-based ontologies focus on the *types* of particulars that exist in reality, RT focusses on the particulars themselves, more concretely on how assertions about the configurations formed between particulars and/or universals should be construed to maximally mimic the structure of reality. Key in RT is (1) the assignment – or reuse in case of former assignment – of instance unique identifiers (IUI) to *every* entity about which some assertion is made, and (2) the use of these identifiers in relational expressions following assertion templates that are maximally self-explanatory and unambiguous [8].

We demonstrate (1) how existing OBO-Foundry ontologies can serve as a source for the representation of all high-level entity types relevant to quality monitoring of devices and data analytic components connected in the IoT for Health, and (2) how RT is expressive enough to represent the relationships enjoyed by instances of these types, and can serve as a basis for analytics regarding the ground truth of assertions.

## 2. Results

Table 1 summarizes the types essential for managing data and metadata to be generated over the IoT for Health with data quality control in mind, specifically the quality aspects *accuracy, consistency and reliability*. Types are universals (U) as introduced in section 1, or defined classes (DC) grouping particulars on the basis of fiat demarcations relevant to some purpose, e.g. to distinguish patients from caregivers [6]. Types are elucidated (E) when primitive or defined (D) in terms of the necessary and sufficient conditions for instantiation. They are taken from BFO [6], the Ontology for General Medical Science (OGMS, [9]), ReMINE's adverse event ontology [10], and the Ontology of Biomedical Investigations (OBI, [11]) or introduced as subtypes from existing types. Further subtyping is possible, but is not relevant for our purposes here.

**Table 1.** Universals (U) and Defined Classes (DC) assessed essential for reporting and analyzing data and metadata generated over the IoT for Health. Terms used in a strict technical sense are formatted in SMALL CAPS and are described either elsewhere in this table (printed in bold) or in the cited reference.

| Type | | Definition (D) or Elucidation (E) |
|---|---|---|
| ASSAY | U | (E) planned PROCESS to produce information about a MATERIAL ENTITY by physically examining it or its proxies [11] |
| BODILY FEATURE | DC | (D) BODILY COMPONENT, BODILY QUALITY, or BODILY PROCESS. [9] |
| CAREGIVER | DC | (D) HUMAN BEING in which there inheres a CAREGIVER ROLE |
| DEVICE | U | (E) OBJECT which manifests causal unity via engineered assembly of components & of a type instances of which are maximal relative to this criterion of causal unity. [6] |
| INTERPRETIVE PROCESS | U | (D) COGNITIVE PROCESS (in brains or through software implementations) which brings into being, sustains or destroys COGNITIVE REPRESENTATIONS on the basis of an **OBSERVATION** [10] |
| IOT FOR HEALTH | DC | (D) OBJECT AGGREGATE which is part of the IoT and is composed out of **DEVICES** and other OBJECTS that generate or analyze **OBSERVATIONS** within a community of **SUBJECTS OF CARE**. |
| SENSOR DEVICE | DC | (D) **DEVICE** in which inheres the FUNCTIONS to perform **ASSAYS** and to generate **OBSERVATIONS** |
| SITE | U | (E) 3-dimensional IMMATERIAL ENTITY that is bounded by a MATERIAL ENTITY or is a 3-dimensional immaterial part thereof. [6] |
| SUBJECT OF CARE | DC | (D) HUMAN BEING undergoing ACTS OF CARE [10] |
| OBSERVATION | DC | (D) **REPRESENTATION** resulting from an ASSAY [10] |
| REPRESENTATION | DC | (D) QUALITY which is about or is intended to be about a PORTION OF REALITY [12] |

Table 2 lists just a few RT statements describing part of a portion of reality evolving over a temporal period *t* during which an inpatient (IUI #1), born at time *t1*, staying in the hospital wearing an RFID tag (#2) since *t2*, is clinically examined in an exam room (#3). The room has an RFID sensor (#4) which is connected to the hospital's IoT for health (#5) and which generated a representation (#109) of the location of the patient's tag when #1 entered room #3 at *t3*. A nurse (#6) measures (#7) at *t4* the patient's temperature (#8) with her personal digital thermometer (#9) which is also connected to #5 since *t5*. has a fingerprint reader to identify patients and a built-in RFID tag to locate its position in the building so that when at *t6* the nurse entered room #3 with the thermometer, sensor #4 generated a representation of its location (#117). When the patient touches the fingerprint reader of the thermometer, a picture (#10) of the patient's fingerprint pattern (#11) is transmitted at *t7* to a fingerprint analyser. This analyser determines (#12) on the basis of another picture of #11 already on file that #10 is about #11, as a result of which it sends the patient's IUI, i.e. '#1', to the thermometer. After the thermometer has registered a value for #10 at *t8*, a representation (#130) is generated asserting that the patient has a temperature of 37°C.

**Table 2.** Some RT statements, preceded by their own IUI, representing part of a scenario of taking a patient's temperature in a healthcare facility with an IoT for Health. The TYPES are listed in Table 1, the relations are defined in (or definable from) the corresponding ontologies, and the temporal operators (e.g. since, includes, during, …) defined in European Norm 12381: Time Standards for Healthcare Specific Problems [13].

| #100: | #1 | *instanceOf* | HUMAN BEING | *since* t1 | #101: | #1 | *instanceOf* | SUBJECTOFCARE | *since* t2 |
|---|---|---|---|---|---|---|---|---|---|
| #102: | #2 | *instanceOf* | DEVICE | *includes* t2 | #103: | #2 | *locatedOn* | #1 | *since* t2 |
| #104: | #4 | *instanceOf* | SENSOR DEVICE | *includes* t | #105: | #4 | *locatedIn* | #3 | *includes* t |
| #106: | #3 | *instanceOf* | SITE | *includes* t | #107: | #4 | *partOf* | #5 | *includes* t |
| #108: | #5 | *instanceOf* | IOT FOR HEALTH | *includes* t | #109: | #2 | *locatedIn* | #3 | *since* t3 |
| #110: | #4 | *authorOf* | #109 | *at* t3 | #111: | #6 | *instanceOf* | CAREGIVER | *includes* t |
| #112: | #7 | *instanceOf* | ASSAY | | #113: | #1..*specifiedInputOf* | | #7 | *during* t4 |
| #114: | #6 | *participantOf* | #7 | *during* t4 | #115: | #9 | *participantOf* | #7 | *during* t4 |
| #116: | #9 | *instanceOf* | SENSOR DEVICE | *includes* t4 | #117: | #9 | *locatedIn* | #3 | *at* t6 |
| #118: | #9 | *partOf* | #5 | *since* t5 | #119: | #9 | *locatedOn* | #6 | *includes* t4 |
| #120: | #4 | *authorOf* | #117 | *at* t6 | #121: | #11 | *instanceOf* | BODILYQUALITY | *since* t1 |
| #122: | #11..*inheresIn* | | #1 | *since* t1 | #123: | #10 | *isAbout* | #11 | *since* t4 |
| #124: | #12 | *instanceOf* | ASSAY | | #125: | #123 | *SpecifiedOutputOf* | #12 | *since* t4 |
| #126: | #11 | *specifiedInputOf* | #12 | *during* t2 | #127: | #8 | *inheresIn* | #1 | *since* t1 |
| #128: | #8 | *instanceOf* | BODILY QUALITY | *since* t1 | #129: | #9 | *authorOf* | #130 | *since* t4 |
| #130: | #8 | *instanceOf* | 37°C | *at* t8 | #130: | #10 | *instanceOf* | OBSERVATION | *since* t7 |

## 3. Discussion

Representations of the sort exhibited in Tables 1 and 2, covering the totality of devices available within an IoT for Health rather than just in the partial scenario developed here, offer ample explicit information to feed algorithms for data quality monitoring by exploiting two specific features of the data collection methodology.

The first one is the multitude of sensor devices that can be used to monitor individual particulars from different perspectives. In the scenario sketched, it is both the RFID tag (#2) of the patient and the fingerprint reader in the thermometer (#9) that provide enough evidence to conclude that it is indeed patient #1 who is examined in room #3. Assertions #103 and #109 together, in combination with the axioms of the ontologies from which the *locatedIn* and *locatedOn* relationships are taken, provide an argument that #1 is in the room, although it might be the case that after #103 was asserted by the reception clerk who gave the RFID tag to patient #1, the patient lost it and picked up another one, or that the clerk made a typo. Similarly, assertions #117 together with #122 through #126 provide evidence for #1 being in the room. But if either something went wrong with the fingerprint analysis or with the RFID tag, both collections of assertions would not lead to the same conclusion what would be a trigger for further verification. The second feature, not worked out in detail in Table 2, is that the patient data can be used to monitor the proper functioning of the IoT devices. If the same scenario applied to several patients would lead to inconsistencies, then it is very likely that either sensor #4 or the thermometer are malfunctioning. This feature makes it clear that what is typically considered metadata, are actually data in their own right.

Although there is no shortage on papers that discuss security and confidentiality risks associated with the IoT for Health, the issue of data quality and anomaly detection is more scarcely dealt with as witnessed by a recent review [14]. An exception is [15] in which a mathematical model towards the reliability of sensor data is proposed, however it's only applicable to continuous sensors with high refresh rate measuring characteristics of ongoing processes (heart beats, continuous blood pressure monitoring) rather than between discrete events. Several papers discuss the potential

use of ontologies in the IoT for Health, but here also mainly for security, e.g. [16]. A literature review over biomedical research papers published between 2001 and 2011 revealed an increasing amount of work on ontology, but little on ontological approaches to data quality [17]. The approach has two limitations. One is the use of ontological realism which is reported to be hard to understand [18]. The other one is the development of not only efficient, but also useful reasoners. Whereas the former requires more education, the latter is a matter of further research and development, including the design of an action logic for inconsistency detection and alerting.

## Acknowledgement

## References

[1] Lugovkina, T. and B. Richards, Clinical events classification for using the EHR to provide better patient care. Stud Health Technol Inform, 2010. 156: p. 167-70.

[2] El Fadly, A., et al., The REUSE project: EHR as single datasource for biomedical research. Stud Health Technol Inform, 2010. 160(Pt 2): p. 1324-8.

[3] Varpio, L., et al., The EHR and building the patient's story: A qualitative investigation of how EHR use obstructs a vital clinical activity. Int J Med Inform, 2015.

[4] Hersh, W.R., et al., Caveats for the use of operational electronic health record data in comparative effectiveness research. Med Care, 2013. 51(8 Suppl 3): p. S30-7.

[5] Sparnon, E. and W.M. Marella, The Role of the Electronic Health Record in Patient Safety Events. Pennsylvania Patient Safety Advisory, 2012. 9(4): p. 113-121.

[6] Arp, R., B. Smith, and A.D. Spear, Building ontologies with basic formal ontology. 2015, The MIT Press,: Cambridge, Massachusetts. p. 1 online resource.

[7] Ceusters, W. and B. Smith, Referent Tracking in Electronic Healthcare Records, in Connecting Medical Informatics and Bio-Informatics. Medical Informatics Europe 2005, R. Engelbrecht, et al., Editors. 2005, IOS Press: Amsterdam. p. 71-76.

[8] Ceusters, W., Chiun Yu Hsu, and B. Smith, Clinical Data Wrangling using Ontological Realism and Referent Tracking. CEUR Workshop Proceedings, 2014. 1237: p. 27-32.

[9] Scheuermann, R.H., W. Ceusters, and B. Smith, Toward an ontological treatment of disease and diagnosis. Summit on Translat Bioinforma, 2009. 2009: p. 116-20.

[10] Ceusters, W., et al., An Evolutionary Approach to Realism-based Adverse Event Representations. Methods of Information in Medicine, 2011. 50(1): p. 62-73.

[11] Jensen, M., et al., Applications of OBI 'assay', in International Conference on Biomedical Ontology. 2014: Houston, TX. p. 96-97.

[12] Smith, B. and W. Ceusters, Aboutness: Towards Foundations for the Information Artifact Ontology, in International Conference on Biomedical Ontology. 2015: Lisbon, Portugal. p. 47-51.

[13] Ceusters, W., et al., TSMI: a CEN/TC251 standard for time specific problems in healthcare informatics and telematics. International Journal of Medical Informatics, 1997. 46(2): p. 87-101.

[14] Rassam, M.A., A. Zainal, and M.A. Maarof, Advancements of data anomaly detection research in wireless sensor networks: a survey and open issues. Sensors (Basel), 2013. 13(8): p. 10087-122.

[15] Marie, P., et al., From Ambient Sensing to IoT-based Context Computing: An Open Framework for End to End QoC Management. Sensors (Basel), 2015. 15(6): p. 14180-206.

# Implementation of a Decision Support System for Interpretation of Laboratory Tests for Patients

Ilya SEMENOV[a] and Georgy KOPANITSA[b,c,1]
[a] *Helix laboratory service, Saint-Petersburg, Russia*
[b] *Institute Cybernetic Center, Tomsk Polytechnic University, Tomsk, Russia*
[c] *Tomsk State University for Architecture and Building, Tomsk, Russia*

**Abstract.** The paper presents the results of the development and implementation of an expert system that automatically generates doctors' letters based on the results of laboratory tests. Medical knowledge is expressed using a first order predicate дцпшс based language. The system was implemented and evaluated in the Helix laboratory service.

**Keywords.** Decision support, Laboratory information system, telemedicine, first order predicates

## Introduction

In Russia many patients address laboratory services directly without a doctor's referral. This causes the problem of interpretation of laboratory test results by the patients who don't have a proper medical background. So the patients require that the laboratory services provide not only the results of the tests but also their interpretations. Automated decision support systems that have proved their efficiency for doctors can be a good solution for this problem [1]. The experience of development and implementation of decision support systems for doctors shows the efficiency of such solutions for the doctors, however, developers face problems when it comes to the decision support for patients. They require different approach in data presentation and interpretation.

The goal of this paper is to present a research and development of a decision support system for the patients of a laboratory service.

To achieve this goal, we have developed a decision support system that solves a classification problem and defines the following parameters based on the results of laboratory tests:

- Diagnosis (group of diagnoses)
- Recommendations to run other laboratory tests
- Recommendation to refer to a specialist doctor

---

[1] Corresponding Author.

## 1. Methods

To achieve the described above goal a decision support system must solve a classification problem by associating a vector of test results to a set of diagnoses and find a set of recommendations associated with every diagnosis from this set.

On the next step we have developed a classification algorithm that has the following possible outcomes:

- Found a set of diagnosis that can be associated to the results of the laboratory test
- No diagnosis found
- Found a set of diagnosis, but the system requires extra test or vital signs to choose the proper diagnosis for this set.

To organize a communication between the system and an expert we have implemented a knowledge representation language (KRL) that is based on the first order predicate logic[6].

After the knowledge representation language was implemented we have developed a graphical user interface to allow experts filling in the knowledge base. For the pilot project we have chosen a limited set of laboratory tests that could be interpreted by the system to test the feasibility of the approach. We have invited 3 laboratory doctors and 3 specialist doctors (gynaecologist, urologist and general practitioner) to fill in the system's knowledge base.

The knowledge representation language, knowledge base and the classification algorithm were developed as a Doctor Ease decision support system, which was implemented in the Helix laboratory service in Saint-Petersburg, Russia. After the system has been implemented we made a qualitative research to evaluate the acceptance of the system among the patients with 100 participants.

## 2. Results

The developed decision support system has a traditional structure and consists of the following modules:

- Data base;
- Data extraction system
- Knowledge base;
- Inference engine;
- Knowledge base editor;
- Explanation system
- Results generator

A structural scheme of the system is presented in the figure 1.
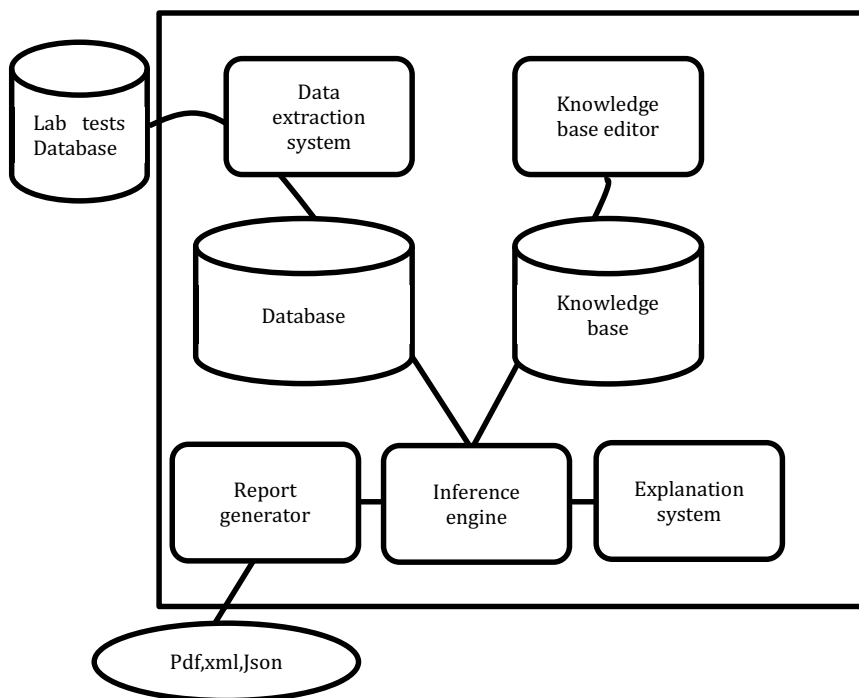
**Figure 1.** Structural scheme of the decision support system.

Each module provides the following functionality to the expert system:

- Data base with a dynamical structure stores facts (test results) and intermediate results of the logical inference. The facts are taken from a laboratory information system (LIS).
- Knowledge base of the DoctorEase stores expert knowledge and inference rules.
- Inference engine applies knowledge and rules from the knowledge base to the facts from the data base to solve the classification task.
- Knowledge base editor provides a user interface to define new knowledge and rules.
- Explanation system analyses the sequence of the rules to explain how the system achieved the result.

The developed decision support system has two main use cases: knowledge acquisition and decision support. Knowledge acquisition mode allows defining inference rules, which are complex objects and each of them adds its element to the resulting inference. The knowledge is defined by associating test results and its reference value to a set of diagnosis. In the decision support mode, the system generates recommendations applying a set of knowledge and rules to the facts that are derived from a LIS data base.

DoctorEase decision support system allows creating queries in the language that is closed to natural. The knowledge representation language is based on the first order logic and the predicates and relationships have meaningful names in Russian so the experts can define knowledge and rules using the terminology they are used to.

## 2.1. Knowledge base organization

The structure of the knowledge base of the system is presented in the figure 2.
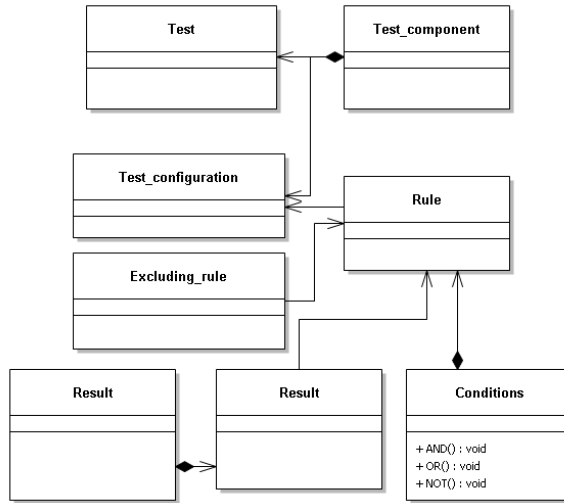


**Figure 2.** Object model of DoctorEase.

On the first step we define a configuration of a laboratory test, which is a complex object consisting of the parameters that are sufficient to make an inference.

- A configuration consists of a laboratory test and inference rules that can be applied to the test.
- A direct rule is an object that is defined for each parameter of a laboratory test along with the conditions for processing these parameters.
- Each rule has a list of exclusion rules, which can exclude direct rules from the inference provided that their conditions are true.
- Laboratory test is a template that consists of laboratory tests' components. For example, a Complete blood count consists of 22 components.
- Laboratory tests are grouped into "orders", which are commercial units that can be ordered by the patients.
- Each rule has a set of conditions that work with comparison operators: =, <>, includes (>= or =<), excludes (>= and =<).
- Conditions are associated with each other by logical operators "and", "or" and "not".

## 2.2. Inference process

After the system has received a notification that the laboratory test results are available it starts the inference according to the following algorithm:

1. Patient's order is analysed to understand if there exist configurations for such orders.
2. Fact (test results) are loaded to the decision support system's data base
3. The inference engine defines a sequence of rules from the knowledge base to be applied to the facts

4. Exclusion rules are applied to the facts to exclude non valid rules from the inference
5. Result blocks are added to the result file according to the rules' sequence.

## 2.3. Implementation

The system was implemented in the Helix laboratory service in Saint-Petersburg, Russia. At the moment it generates about 3500 reports a day. The implementation of the system allowed increasing the number of patients who refer to a doctor after laboratory tests by 14%. A qualitative study with 100 patients demonstrated a high acceptance of the system. The majority (82%) of the patients reported that they trust the system and follow its advice to visit a doctor if necessary.

## 3. Discussion

The paper presents a process of development and implementation of a decision support system for laboratory service patients. The system allows patients reading and understanding medical records in natural language. For the laboratory service the system allowed increasing the level of satisfaction of the patients and the number of patients who came back to the laboratory service for more detailed testing.

Current research is focused on the extension of the knowledge representation language by adding an ability to work with fuzzy sets [7] in definition of knowledge and rules. We also are studying the possibility to validate the reports that are produced by DoctorEase to enable the system acquiring knowledge based on its experience, knowledge and rules. We also are studying the possibility to validate the reports that are produced by DoctorEase to enable the system acquiring knowledge based on its experience.

## References

[1] L. Ahmadian, M. van Engen-Verheul, F. Bakhshi-Raiez, N. Peek, R. Cornet, and N.F. de Keizer, The role of standardized data and terminological systems in computerized clinical decision support systems: literature review and survey, Int J Med Inform 80 (2011), 81-93.
[2] G. Kopanitsa, Standard based multiclient medical data visualization, Stud Health Technol Inform 180 (2012), 199-203.
[3] G. Kopanitsa, Evaluation study for a multi-user oriented medical data visualization method, Stud Health Technol Inform 200 (2014), 158-160.
[4] G. Kopanitsa, Z. Tsvetkova, and H. Veseli, Analysis of metrics for the usability evaluation of EHR management systems, Stud Health Technol Inform 180 (2012), 358-362.
[5] G. Kopanitsa, Z. Tsvetkova, and H. Veseli, Analysis of metrics for the usability evaluation of electronic health record systems, Stud Health Technol Inform 174 (2012), 129-133.
[6] R.S. Michalski, Pattern recognition as rule-guided inductive inference, IEEE Trans Pattern Anal Mach Intell 2 (1980), 349-361.
[7] K. Boegl, K.P. Adlassnig, Y. Hayashi, T.E. Rothenfluh, and H. Leitich, Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system, Artif Intell Med 30 (2004), 1-26.

This page intentionally left blank

# Clinical Information Systems and Data Reuse

This page intentionally left blank

# Exploring Barriers and Opportunities for Adoption of Web Portals in Russia

Georgy KOPANITSA [a,b,1] and Vladimir YAMPOLSKY [b]
[a] *Institute Cybernetic Center, Tomsk Polytechnic University*
[b] *Tomsk State University for Architecture and Building*

**Abstract.** The aim was to study the opinions of tuberculosis patients and doctors and identify perceived opportunities and barriers to using a Web portal to optimize its use. The perceptions of 30 tuberculosis patients and 18 doctors (10 general practitioners and 8 tuberculosis specialists) from Tomsk, Russia were collected through semi-structured interviews. The responses were analysed for content using principles of the grounded theory and thematic analysis, in order to gain insight into the participants' beliefs and attitudes towards adopting tuberculosis web-portal to increase the efficiency of the treatment and rehabilitation process.

**Keywords.** Web-portal, qualitative study, grounded theory, evaluation

## Introduction

Within many healthcare systems especially with many remote areas, there has been a growing interest on the value of telehealth for improving quality, accessibility and cost-effectiveness of care for people with long term complex health and social care needs such as tuberculosis. Such technologies are varied, but a distinction is generally drawn between 'telehealth' and 'telecare'. Telehealth technologies allow remote medical data exchange (e.g. blood pressure, temperature, blood sugar) and additional information between a patient and health care professionals to assist in the diagnosis and management of a health conditions. Telecare allows remote monitoring of changes in patients' conditions or environment [1; 2]. Telehealth and telecare solutions have been implemented and evaluated in various clinical settings. Some evidence of positive outcomes has been reported including improved clinical indicators and reduced health service use. However, a number of studies have demonstrated that these types of solutions often fail to be successfully implemented and adopted within routine healthcare. This is caused by the complexity of telehealth and telecare projects. The use of Web portals has several benefits. It can enhance communication between patient and health care professionals, allow patients to play a more active role in their own treatment and self-management, increase self-efficacy, and patients can feel that other non-acute concerns are valued because of an email function. The use of Web portals shows promising results in treatment outcomes. With the growing number of people with tuberculosis in Russia, the use of patient portals for tuberculosis management becomes more important to cope with the burden on health care. Implementation of such projects involves a major work of a team of technicians, doctors and patients. Different user

---

[1] Corresponding Author.

groups have different expectations on how telemedicine applications can affect their everyday life. This creates one of the major obstacles in adopting of telemedicine solutions. To help telemedicine software developers to overcome these problems we have conducted a qualitative study to analyse what doctors and patients expect from telemedicine solutions on the example of tuberculosis patients' portal. [3-5]. The web portal provides patients an access to their medical health record where they can contribute vital signs to the health care provider. Doctors can access patients' medical data and give recommendations to the patients. The web portal is completely free of charge for the patients.

## 1. Methods

The study was made with semi-structured one-to-one expert interviews based on an interview topic guide [6; 7]. A semi-structured topic guide was drafted by the study team, reviewed by the wider research team and circulated for expert feedback from the health care professionals involved in service delivery. The guide was used flexibly during the interviews to keep the natural flow of conversation and to allow participants to freely discuss their experiences. Interview questions agenda:
1.  What is your opinion about the telehealth portal currently used to communicate with your patients?
    1.1.  What do you think is the best or worst aspect of this tool, and why?
    1.2.  What do you think needs most improvement, and why?
2.  Please state any comments/suggestions that you might have for the portal, if any.
3.  What would you want and need from a telehealth tuberculosis portal?
    The interview topic guide comprised open questions to provide good insights into the participant's opinion, without influencing the speaking flow by open questions was expected to be gained along with new aspects not named before. Participants were recruited from the Tuberculosis treatment and rehabilitation centre in Tomsk, Russia. All the participants have been using the portal for not less than a month by the moment of the study begins. To study expectations and barriers of two user categories we divided the study participants into two groups: doctors and patients with 8 TB specialists and 10 general practitioners, and 30 tuberculosis patients having a rehabilitation course. The interviews were performed by two experts. Written consent was acquired from each interviewee and each interviewee was reminded of their entitlement to withdraw their data from the study database for up to three months after their approval and return of their own interview transcript. The responses of each participant were analysed by employing a qualitative approach in general and particularly by combining thematic analysis with the principles of the grounded theory [8]. Thematic analysis and principles of the grounded theory were employed in the following systematic manner. The coding began by following the process of open-coding, which involves the systematic reading employing a line-by-line analysis considering every comment produced by participants.

## 2. Results

The thematic analysis of the answers has resulted in 5 main high-level core topics related to the participants' responses to the interviews, namely: (1) communication, (2) intention to use, (3) technical competence, (4) responsibility and (5) data protection. The first core

topic highlighted the potential that the better communication between doctors and patients. The second core themes that were identified brought to the fore the participants' intentions to use the portal in their routine life, and that they believed it would be easy for to systematically interact. The third core topic raised the problem of technical competence required to successfully use the tool. The fourth and fifth topics reveal the problems of responsibility and data protection awareness. Tables 1 and 2 present a pool of selected participant responses that support these results.

**Table 1.** Selected patients' (P) responses and core topics

| Core topic | Selected responses |
| --- | --- |
| Communication | I can get treatment recommendations at home, no need to visit a doctor [P1] |
| | I expected that a doctor would give me recommendations every time I submit my data [P2] |
| | I can work with the portal anywhere, no need to go home or visit a doctor if a need a recommendation [P3] |
| | I can ask my relative or friend to help me understand the treatment recommendations given by a doctor [P4] |
| Intention to Use | The portal allows to quickly update the treatment if my health conditions change [P6] |
| | It reminds me to take my medicine in the right time [P7] |
| | I am glad to use these new technologies, I feel like my treatment has come to a new level [P8] |
| Technical competence | I learned how to use it in few minutes [P9] |
| | Very user friendly interface, I feel like I cannot do anything wrong [P10] |
| | The portal has very nice design, this influences my mood in a good way [P11] |
| Responsibility | Using this portal makes me more responsible for my health [P12] |
| | I feel involved in the treatment process [P13] |
| Data protection | I am aware how my medical data is protected  [P14] |
| | Now I cannot lose my lab test papers. I have everything in one place [P15] |
| | I don't know how my data is protected, however I trust my doctors [P16] |

**Table 2.** Selected doctors' (D) responses and core topics

| Core topic | Selected responses |
| --- | --- |
| Communication | It  allows me to see the data in a very usable form: diagrams, tables.[D1] |
| | Patients still expect me to be available online when they have submitted their data, even if it is day off or out of office hours [D2] |
| | I can react on acute situation and give my recommendations to the patients quickly [D3] |
| | I still have to do a lot of paper work, my work has almost doubled [D4] |
| | Patients think that I must react event if their health status is normal [D5] |
| Intention to Use | The health status of my patients has increased [D6] |
| | My patients started taking their medicine on a regular basis [D7] |
| | I can change the treatment plan without visiting a patient [D8] |
| | I would use the portal instead of all paper work I have to do [D9] |
| Technical competence | The portal is very easy to use [D10] |
| | The tool does not require advanced technical competence and much easier than previous EHR systems [D11] |
| | I can have data of all my patients on one screen [D12] |
| | The portal allows to have all necessary data on the screen, no need to search previous records of a patient [D13] |
| Responsibility | I cannot be sure that the patients always read my recommendation in time. But I am still responsible for their health status, the portal cannot replays a visit in every situation [D14] |
| | I don't know if the electronic record is a valid medical document [D15] |
| | I feel more responsibility for the patients if I know they are involved in the process [D16] |
| Data protection | I am aware that the medical data is protected, because otherwise I can have law troubles [D17] |
| | I follow all the data protection instructions to support patients' privacy [D18] |
| | My password is too long due to data protection, but I can handle it [D19] |

These core topics revealed that participants from different groups have different expectation from the tool.

- **Communication:** Despite the fact that both doctors and patient were particularly enthusiastic with the ability to communicate remotely [table 1, P1, P3, P5; table 2, D4] they perceive this ability in a different manner. The patients expected a 24/7 direct communication with a doctor [P2]. Most of them claimed that they expect a doctor's reaction after every data exchange fact. However, doctors reported that they expected that the tool would improve their communication with patient, but won't cause them more work to do [D2, D4, D5]. Specifically, doctors [D3] explained that the tool is very useful in acute or changing conditions, but in general it causes much work load due to the fact they need to duplicate their recommendations on paper [D4] and they get many requests from the patients even in the situation when a doctor's interference is not needed [D2, D5].

- **Intention to Use:** The patients also mentioned the clarity [P4] of the recommendations and possibility to ask their friend or relatives to help them understanding the recommendations [P5]. Many patients discussed their decision to use the portal because they feel they get better service [P6-P7]. The patients reflected on the place of technology in the modern world and their intention regarding the use of contemporary communication technologies They feel that they use the most modern technologies, which inspires them to follow all the doctors'recommendations [P8]. As for the doctors, they see a practical value in the portal. They have an intention to use it instead of traditional paperwork [D9] and to save time visiting a patient [D8]. They have also mentioned that the patients have started following their recommendations and these has led to increase of the patients'health status [D6-D7].

- **Technical competence:** The overall responses indicated that the majority of the participants did not mention any difficulties with using the portal [D10-D11, P9-P10]. Doctors have mentioned that the information visualization and reporting services of the portal allow them having all relevant data on the screen and provide very efficient searching facilities [D12 - D13]. Patients have also mentioned the design of the portal that is very appealing and increases users' experience [P11]. It must be remarked that from the transcribed responses, the general trend was that participants were enthusiastic about the functionality of the portal and majority of the participants mentioned a very good training that they had had before they started using the portal.

- **Responsibility:** The patients indicated that the using of the portal makes them feel involved in the treatment process and feel more responsibility for their health status [P12-P13]. The doctors have also mentioned this change of patients' attitude and behaviour [D16], however, the doctors feel a little insecure about the legal validity of the recommendations given through the portal [D15]. They worry that the patients will read their recommendation in time and understand them properly. Doctors have also mentioned that they are still responsible for the patients and would not like the portal to replace all the patients' visits [D14].

- **Data protection:** The overall data protection awareness was observed in the majority of the interviews [D17-D19], [P14-P16]. Users understand the importance of medical data protection and are ready to follow the data protection instructions even if they are not easy to follow [D19]. The patients see the other aspect of data protection: they feel that if data is stored electronically it is unlikely to get lost in comparison with the paper records that are being lost by the clinics on a regular basis.

## 3. Discussion

The study has revealed differences between doctors and patient's expectations of the web portal. Doctors have reported two major explications concerning time saving of the anamnesis collection processes as some data is already submitted by the patients and being able to quickly react to the acute problems of the patients. Patients have reported that they expect that the portal will give them 24/7 contact with a doctor. They expect a reaction form a doctor as soon as the data is submitted and the patients think that something might be wrong. This conflict of expectations has become the major obstacle to the successful implementation of the project as some doctor's decline to contact patients after each data transfer event and the patients don't want to submit data without a doctors' reaction.

## 4. Conclusion

As the need for quality health care provision continues to expand, web portal technologies can be an efficient and intuitive approach for improving medical practices. The current paper explores from a qualitative perspective on the attitudes of 30 patients and 18 doctors towards using a tuberculosis web portal to report patients' vital signs and produce treatment recommendations. The findings from our study in the area of tuberculosis treatment and rehabilitation suggested that the participants generally accepted the introduction of web-portal for reporting their health status and getting recommendation from the doctors as an alternative to the traditional doctor visits.

## References

[1]   H.F. Chen, M.C. Kalish, and J.A. Pagan, Telehealth and hospitalizations for Medicare home healthcare patients, Am J Manag Care 17 (2011), e224-230.
[2]   N.R. Chumbler, D. Haggstrom, and J.J. Saleem, Implementation of health information technology in Veterans Health Administration to support transformational change: telehealth and personal health records, Med Care 49 Suppl (2011), S36-42.
[3]   M. Brown and N. Shaw, Evaluation practices of a major Canadian telehealth provider: lessons and future directions for the field, Telemed J E Health 14 (2008), 769-774.
[4]   E. Dillon and J. Loermans, Telehealth in Western Australia: the challenge of evaluation, J Telemed Telecare 9 Suppl 2 (2003), S15-19.
[5]   I. Ellis, Is telehealth the right tool for remote communities? Improving health status in rural Australia, Contemp Nurse 16 (2004), 163-168.
[6]   A. Ostenfeld-Rosenthal and H. Johannessen, How to capture patients' concerns and related changes: comparing the MYCaW questionnaire, semi-structured interview and a priority list of outcome areas, Complement Ther Med 22 (2014), 690-700.
[7]   J.H. Isaksen, N.T. Hertel, and N.K. Kjaer, Semi-structured interview is a reliable and feasible tool for selection of doctors for general practice specialist training, Dan Med J 60 (2013), A4692.
[8]   Grounded Theory: A Practical Guide Birks Melanie and Mills Jane Grounded Theory: A Practical Guide 208 pages pound26.99 Sage 9781446295786 1446295788, Nurs Stand 30 (2015), 28.

# Elderly Surgical Patients: Automated Computation of Healthcare Quality Indicators by Data Reuse of EHR

Grégoire FICHEUR[a,b], Aurélien SCHAFFAR[d], Alexandre CARON[b], Thibaut BALCAEN[b], Jean-Baptiste BEUSCART[a,c], and Emmanuel CHAZARD[a,b,1]

[a] *Univ. Lille, EA 2694 Public Health, F-59000 Lille, France*
[b] *CHU Lille, Public health department, F-59000 Lille, France*
[c] *CHU Lille, Geriatrics department, F-59000 Lille, France*
[d] *Department of Medical Information, St Philibert Hospital, GHICL, Lille, France*

**Abstract.** The objective of the work is to implement and evaluate the automated computation of 9 healthcare quality indicators, by data reuse of electronic health records, in the field of elderly surgical patients. Methods: Data are extracted from EHR, including administrative data, ICD10 diagnoses, laboratory results, procedures, administered drugs, and free-text letters. The indicators are implemented by a medical data reuse specialist. The conformity rate is automatically computed (3.5 minutes for 15,000 inpatient stays and 9 indicators). A medical expert reviews 45 stays per indicator. The precision is the proportion of non-conform inpatient stays among the cases detected as non-conform by the algorithms. Results: the paper describes the implemented algorithms, the conformity rates and the precisions. Two indicators have a precision of 0%, 3 indicators have a precision of 40 to 60%, and four indicators have a precision from 80 to 100% (for 2 of them, the conformity rate is lower than 2.5%!). This demonstrates that automated quality screening is possible and enables to detect threatening situations. The implementation of the indicators requires special skills in medicine, medical information sciences, and algorithmics. Failures of precision are mainly due to defaults of information quality (missing codes), and could benefit from text analysis.

**Keywords.** Quality Indicators, Process Assessment, Guideline Adherence, Data reuse, Electronic Health Records, Geriatrics

## Introduction

The first step to improve the quality of care is to measure it by mean of quality indicators. Among them are process indicators, which measure the percentage of adherence to a guideline. They are widely used in healthcare. Contrary to input indicators, they are closer to patients' outcome [1]. Contrary to outcome indicators, they are not impacted by the initial severity of the patients [2–6], and are then easier to interpret [7,8]. Finally, they directly enable to identify areas of improvement [2,3].

In a previous work, a systematic literature review [9] was performed. It enabled to identify 8,744 papers, among which 126 papers described 440 process indicators. Some

---

[1] Corresponding Author.

of those indicators (22.3%) could be automatically computed using a minimal dataset including diagnoses, drug administrations, medical procedures, administrative data, laboratory results, free-text reports with basic keyword research, linkage with the patient's previous stays, and dependence assessment.

The objective of this work is to implement and evaluate process indicators. To provide the reader with a consistent set of indicators, 9 indicators about surgery on elderly are selected from a unique scientific paper published by McGory et al. [10].

## 1. Methods

A database of 15,000 randomly-selected and de-identified inpatient hospital stays is extracted from a French community hospital. It notably contains administrative data (patient flow, age, gender, etc.), diagnoses (ICD10 codes without date), procedures (French CCAM codes with dates), administered drugs (ATC codes with dates, routes and doses), laboratory results (with dates, values and units), and free-text letters.

The indicators published by McGory et al. [10] are available as free-text. A physician, with a personal background of computer scientist, implements them as a set of PHP code in command line interpreter mode. Each indicator is defined as a function, composed of inclusion criteria, and conformity criteria. For each stay, the function returns "NA" if the stay is not included, "1" if the stay complies with the guideline, and "0" in other cases (Figure 1). The conformity rate is the average of the return values.
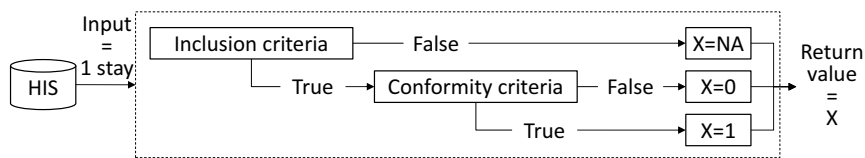


**Figure 1.** Each function is composed of inclusion criteria, and conformity criteria.

The implementation process takes into account the available data, and aims at increasing the precision of the detection of non-compliant cases. Over-alerting is decreased by relaxing the conformity criteria, or by tightening the inclusion criteria (respectively arrows (1) and (2) on Figure 2). The purpose is to identify fields requiring a priority intervention of a quality management team, without demotivating caregivers.

Finally, for each rule, an expert reviews randomly-selected cases: 15 excluded cases (but with interesting keywords in the discharge letter), 15 non-conform cases, and 15 conform cases (if possible). The expert reads completely the free-text documents.



**Figure 2.** Contingency table of the automated output and the expert advice: conformity rate and precision.

## 2. Results

The total processing time is 3.5 minutes for 9 indicators and 15,000 inpatient stays using a personal computer. This section is composed of one subsection for each indicator. "*n*" denotes the number of included stays, "*conformity*" is the automatically computed conformity rate, and "*precision*" is the ability of the program to detect non-conform cases according to the experts. Important changes in criteria are commented, as well as the lack of precision. From this point, "*If (...)*" stands for "*If an elderly patient is undergoing surgery*". Original formulations of indicators are given in [10].

**Indicator #01: creatinine clearance dosage**
Formulation: If (…), then creatinine clearance should be estimated.
Inclusion criteria: [age≥75 & surgical stay & emergency admission]. As it is not possible to screen ambulatory laboratory exams in case of pre-planned admission, only emergency admissions are analysed.
Conformity criteria: [creatinine clearance or creatinine dosage before surgery]. We assumed some physicians were able to compute the creatinine clearance by hand from the creatinine dosage.
Results: n=238, conformity=86.6% [81.6-90.7], **precision=86.7%** [59.5-98.3]. Some dosages could be absent from databases but cited in the discharge letter.

**Indicator #02: confusion**
Formulation: If (…) and has a new diagnosis of delirium, then an evaluation should be undertaken for: infection, electrolyte abnormalities, hypoxia, uncontrolled pain, urinary retention or faecal impaction, use of sedative-hypnotic drugs.
Inclusion criteria: [age≥75 & surgical stay & ICD code (F05* | R410*)]. It was not possible to analyse the date of diagnosis.
Conformity criteria: [postoperative dosage of $Na^+$ & $K^+$ & Creat. & Uremia & Glycemia]. It was not possible to automatically detect whether the physicians had searched for other clinical factors or drugs administrations.
Results: n=8, conformity=75 % [34.9-96.8], **precision=0%.** Capillary blood glucose had been measured from patient bedside but not traced in the databases. In 2 cases there was a post-operative confusion, without ICD10 code.

**Indicator #03: Beta-blockers**
Formulation: If (…) and takes a beta blocker as an outpatient, then beta blocker therapy should be continued postoperatively.
Inclusion criteria: [age≥75 & surgical stay & ATC code C07* at d0 or d1 & no death]. The outpatient treatment was not available. The contraindications could not be traced using ICD10 codes. Dead patients were excluded.
Conformity criteria: [ATC code C07* administered from d2 to day prior discharge]
Results: n=18, conformity=27.8% [9.7-53.5], **precision=46.2%** [19.2-74.9]. All false positives were due to some missing ATC codes in the database.

**Indicator #04: intravenous antibiotic prophylaxis (onset)**
Formulation: If (…), then intravenous antibiotic prophylaxis should be started within 1h of skin incision.
Inclusion criteria: [age≥75 & surgical stay & no ICD code of (bacterial infection, unprecise infection, or open fracture)]. We excluded cases were an oral antibiotic treatment was already indicated.
Conformity criteria: [intravenous administration of antibiotic the day of surgery]

Results: n=607, conformity=2.4% [1.5-4.2], **precision=93.3%** [68.1-99.8]. All false
positives were due to the absence of ICD10 codes despite an active infection, for
which the oral antibiotic treatment was appropriate.

**Indicator #05: oral antibiotic prophylaxis (continuation)**
Formulation: If (…), intravenous antibiotic prophylaxis should be discontinued within
24h after surgery (48h for cardiac surgery).
Inclusion criteria: [age≥75 & surgical stay & no ICD code of (bacterial infection,
unprecise infection, or open fracture)]. We excluded cases were an antibiotic
treatment was necessary.
Conformity criteria: [no ATC code J01* at d3 of surgery]. The type of surgery was not
taken into account, and a permissive delay was used.
Results: n=607, conformity=2.3% [1.3-3.9], **precision=93.3% [68.1-99.8].** All false
positives were due to the absence of ICD10 codes despite an active infection.

**Indicator #06: deep venous thrombosis prevention**
Formulation: If (…), then deep venous thrombosis prophylaxis should be provided
(unfractioned or low molecular weight heparin) or document why not appropriate.
For cancer or previous thromboembolism, mechanical prophylaxis should be added.
Inclusion criteria: [age≥75 & surgical stay & CCAM code of (major surgery of lower
limbs | major digestive surgery) & no ICD10 code of heparin contra-indication]. It
was not possible to trace mechanical prophylaxis. Therefore, we traced heparin
administration, only for high risk surgeries. We searched for contra-indication.
Conformity criteria: [ATC code B01AB*]
Results: n=314, conformity=70.0% [65.6-76.0], **precision=53.3% [26.6-78.7].** False
positives are mainly due to the inclusion of non-surgical inpatient stays.

**Indicator #07: anemia treatment**
Formulation: If (…) and has anemia, then the following should be set up prior to
surgery: iron, vitamin C, erythropoietin, blood transfusion if hemoglobin < 7 g/dl.
Inclusion criteria: [age≥75 & surgical stay & hemoglobin value<10]. The threshold was
not specified. The severity of anemia is handled in indicator #08.
Conformity criteria: [ATC code B03A* or B05AX01 | ICD10 code Z5130 |
transfusion]
Results: n=217, conformity=60.4% [53.5-66.9], **precision=86.7% [59.5-98.3].** False
positives are due to the absence of administered drugs from the databases.

**Indicator #08: anemia transfusion**
Formulation: If (…), unless otherwise contraindicated or refused by the patient, then
he/she should receive blood transfusion at the following hemoglobin/hematocrit
threshold: 8/24 (man), 7/21 (woman).
Inclusion criteria: [age≥75 & surgical stay & ((man & (hemoglobin<8 |
hematocrit<24)) | (woman & (hemoglobin<7 | hematocrit<21)))]
Conformity criteria: [ATC code B05AX01 | ICD10 code Z5130 | CCAM transfusion]
Results: n=19, conformity=73.7% [48.8-90.9], **precision=60.0% [14.7-94.7].** False
positives are mainly due to the absence of transfusion encoding.

**Indicator #09: postoperative fever**
Formulation: If (…) and has a new fever, the following should be performed: urinalysis
and urine culture, wound examination, blood cultures of central venous line or
catheter, chest radiograph, blood culture.

<u>Inclusion criteria:</u> [age≥75 & surgical stay & ICD10 code (R50* | R65*)]. Note that if the etiology of a fever is found, then it is encoded and the fever is not encoded.

<u>Conformity criteria:</u> [CCAM code of chest radiograph]. Clinical exams and some laboratory exams were not available.

<u>Results:</u> n=4, conformity=75% [19.4-99.4], **precision=0.0%.** The inflammatory syndrome was present before the surgery (no precise date).


## 3. Discussion

The automated computation of process quality indicators is feasible, and enables to quickly identify threatening situations due to failures in guideline adherence. However, even indicators that have been published require a medical interpretation of free-text, and deep modifications of their algorithms. Those modifications, as well as the implementation, require special skills in medicine, medical information sciences, and algorithmics in the same time. To our knowledge, rule management systems would bring no help for that important and expert-based task. In addition, the algorithms are complex, polymorphic and would however require hard code writing.

However, some indicators suffer from problems in data quality, such as missing codes. Their lack of precision could be improved by free-text analysis. If such indicators are deployed, their precision should be re-evaluated in each new setting, because quality failures may be hospital-dependent. Finally, those indicators could be used to identify clinical cases to support morbidity and mortality reviews.


## References

[1] Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD Health Care Quality Indicators Project. *Int J Qual Health Care* **18** Suppl 1 (2006),5–13.

[2] Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual Health Care* **13** (2001), 475–80.

[3] Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. *Int J Qual Health Care* **13** (2001), 469–74.

[4] Lingsma HF, Dippel DWJ, Hoeks SE, Steyerberg EW, Franke CL, van Oostenbrugge RJ, et al. Variation between hospitals in patient outcome after stroke is only partly explained by differences in quality of care: results from the Netherlands Stroke Survey. *J Neurol Neurosurg Psychiatry* **79** (2008), 888–94.

[5] Nicholas LH, Osborne NH, Birkmeyer JD, Dimick JB. Hospital process compliance and surgical outcomes in medicare beneficiaries. *Arch Surg* **145** (2010) 999–1004.

[6] Ploeg AJ, Flu HC, Lardenoye JHP, Hamming JF, Breslau PJ. Assessing the quality of surgical care in vascular surgery; moving from outcome towards structural and process measures. *Eur J Vasc Endovasc Surg* **40** (2010), 696–707.

[7] Kerr EA, Smith DM, Hogan MM, Hofer TP, Krein SL, Bermann M, et al. Building a better quality measure: are some patients with "poor quality" actually getting good care? *Med Care* **41** (2003), 1173–82.

[8] Powell AE, Davies HTO, Thomson RG. Using routine comparative data to assess the quality of health care: understanding and avoiding common pitfalls. *Qual Saf Health Care* **12** (2003), 122–8.

[9] Chazard E, Babaousmail D, Schaffar A, Ficheur G, Beuscart R. Process assessment by automated computation of healthcare quality indicators in hospital electronic health records: a systematic review of indicators. *Stud Health Technol Inform* **210** (2015), 867–71.

[10] McGory ML, Kao KK, Shekelle PG, Rubenstein LZ, Leonardi MJ, Parikh JA, et al. Developing quality indicators for elderly surgical patients. *Ann Surg* **250** (2009), 338–47.

# Clinical Data Models at University Hospitals of Geneva

Dina VISHNYAKOVA [a,1], Christophe GAUDET-BLAVIGNAC [a], Philippe
BAUMANN [a], and Christian LOVIS [a,b]

[a] *Division of Medical Information Sciences, University Hospitals of Geneva, Geneva,*
*Switzerland*
[b] *University of Geneva, Geneva, Switzerland*

**Abstract.** In order to reuse data for clinical research it is then necessary to overcome
two main challenges – to formalize data sources and to increase the portability. Once
the challenge is resolved, it then will allow research applications to reuse clinical
data. In this paper, three data models such as entity-attribute-value, ontological and
data-driven are described. Their further implementation at University Hospitals of
Geneva (HUG) in the data integration methodologies for operational healthcare data
sources of the European projects such as DebugIT and EHR4CR and national project
the Swiss Transplant Cohort Study are explained. In these methodologies the clinical
data are either aligned according to standardised terminologies using different
processing techniques or transformed and loaded directly to data models. Then these
models are compared and discussed based on the quality criteria. The comparison
shows that the described data models are strongly dependent on the objectives of the
projects.

**Keywords.** Clinical data model, Data Integration, Data Ontology, EHR

## Introduction

The research of eHealth usually faces the problem of data reuse and consequently the
semantic data interoperability. Solution of this problem is essential for the sustainable
use of information. Electronic health records (EHR) are the growing part of the eHealth,
ranging from clinical findings to genome structures where continuous electronic
processes improve coordination and rapid exchange of information among stakeholders.
However, the up cycling (secondary utilisation of data) of clinical data to improve
healthcare quality and patient safety are very limited. Therefore, the need to integrate
heterogeneous data from multiple sources and sharing information in a distributed and
collaborative environment are highly challenging. The data integration process for the
research environment is a complex task, which has to take into account the following
questions: 1) How the system will be used? 2) Who will use the system?

The data model could influence the research facilities starting from defining what
kind of data can be stored to how the information will be queried and extracted. To date,
there are some models, which became widely used in research domains, for instance
OMOP Data Model of OHDSI [1], archetypes of openEHR [2] and the dimensional
modelling of i2b2. Since there is no solution to identify the generic data model able to

---

[1] Corresponding Author.

be efficient for any research need we propose various data models selected for the projects conducted at University Hospitals of Geneva: 1) the cohort project aims to acquire high quality data about patients with chronic diseases. The Swiss Transplant Cohort Study (STCS), launched in 2008, was the first module implemented in this project. This module is daily used in Switzerland to follow the complete population of transplanted patients since 2008. Other active cohorts have since joined the project to allow follow-up and statistical studies over more than 5000 patients; 2) In a context where the emergence and increase of antimicrobial resistance is problematic, the European FP7 DebugIT project [3], [4] aimed to improve and monitor prescriptions of antibiotics and thus reduce antimicrobial resistance; 3) EHR4CR - provides adaptable, reusable and scalable solutions for data reuse systems of EHRs for clinical research. The project addresses four main scenarios: the feasibility of clinical protocols, identification and patient recruitment, execution of clinical trials and reporting of side effects [5].

We should notice that this paper doesn't aim the description of the legal dimension of medical data reuse.

## 1. Methods

We present the data integration methodologies implemented by University Hospitals of Geneva (HUG), which promote technical and semantic interoperability for operational healthcare data sources. The selection of the data model is based on the needs of the projects. We present 3 different data models: 1) Entity-attribute-value evolved from the legacy system 2) ontology-driven data model which is built from scratch in order to extract antibiotic-relevant information and 3) data-driven model, the existing data model (i2b2 star schema) is adapted to the needs of the project.

### 1.1. Entity-attribute-value model

The STCS project and as well as the 4 other cohort projects of HUG are based on the entity-attribute-value (EAV) information model. Since at the beginning it was an institutional project the information model of the project is based on the legacy clinical system. Every row in EAV model is composed of three fields: 1) an entity representing a described item (a consultation with a patient); 2) an attribute describing the entity (e.g. cardiac frequency) and 3) the value of the attribute (e.g. 64 beats/min), which can be of any type. For each cohort project an ad-hoc web application was developed to allow an interaction with the clinical data. This interaction is based on the object-relational mapping tool. The data-model is patient oriented. The semantics of data is based on the terms, specially developed by clinicians to match the needs of the project. Since the number of patients per institution was not large, the population of data model is done manually by data-managers of the defined medical centres.

The data is usually extracted from EAV model on demand and sent to a central-data-manager, it is then cleaned and processed to be easily usable for analysis. There are currently more than 80 research projects that have passed the ethical and scientific committees and use these data for research.

## 1.2. Ontological model

The DebugIT project performed a bottom-up approach for the data integration. The data model is done from scratch. Initially, clinical data from different hospitals were collected and organized in virtualized clinical data directories. Thus, each pilot site has a relational database fulfilled by information on microbiology, medicaments and patient administration. This information was standardised by terminologies such as NEWT and WHO-ATC. The goal of a pilot site – to formalise the data so that it can be ubiquitously accessed using a formal query language. The queries constructed in a manner to answer a specific question related to antibiotic resistance, e.g. "What is the evolution of <u>bacteria</u> resistance to <u>antibiotic</u> during <u>period</u> at <u>location</u>?" For this purpose, the underlying database content is transformed into a formal language representation. This is achieved by defining the elements, their classes, properties, instances and relationships, using a formal ontology language [6]. In this case, the formalized data model (FDM) is a direct map of the original database schema to an ontological model. The FDM vocabulary directly reflects the table and column names of the source model. The FDM is then connected to the underlying non-formal database so that it can provide data in the RDF format and be accessed through a SPARQL query protocol.

## 1.3. Data-driven model

In the context of EHR4CR project, the HUG as a pilot site has chosen a model-driven approach for data integration. The pilot sites agreed to expose their data in a form of Clinical Data Warehouse (CDW). HUG's CDW is based on the platform of Informatics for Integrating Biology and the Bedside (i2b2) [7]. Since the scenarios of the project are patient-centric. The choice of i2b2 data model is rather rational, it fits the patient-centric scenarios and is represented as a five-axis star: Patient, Visit, Observation, Concept and Provider [7]. The clinical data of HUG corresponding to such axis as Patient, Visit and Provider were integrated directly to the i2b2 schema. Since clinical data of laboratory analysis (Labs) is not using the international standardization code such as LOINC we have created local concepts of Labs for the CONCEPT dimension. The semantic interoperability with the research environment is implemented outside of the data model. In order to federate all heterogeneous clinical data sources the project had addressed some semantic interoperability aspects and the clinical data storage model through: 1) the terminology mapping service for dynamical translation of concepts of the central to local terminologies [8]; 2) the query system developed to transform eligibility criteria into queries that interrogate heterogeneous local data warehouses [9] .

## 2. Results

In order to compare the three different data models we have used the data quality criteria defined in [10], see Table 1. In this table the criteria defined in [10],  is adapted for the clinical data models as following: 1) completeness – does the data coverage fit the project needs ; 2) integration – does the model link all data dimensions (Patient, observation, visit, laboratory analysis) correctly? ; 3) understandability – do the data structure and concept make sense to all end users (data investigators, data managers, clinicians, etc)? ; 4) simplicity – is it easy to transform data elements to the model? 5) flexibility – is it

possible to extend the project scope within the data model (e.g. to add new elements without changing the schema)?

**Table 1.** Evaluation of data models according to quality dimensions

| Dimension | Ontology-driven | Data-driven | EAV |
|---|---|---|---|
| Completeness | yes | yes | yes |
| Integration | yes | yes | yes |
| Understandability | partly | yes | yes |
| Simplicity | no | yes | no |
| Flexibility | no | no | yes |

Since the models were maintained in the framework of the projects we have excluded the criterion as *implementability* from the list of dimensions. This criterion depends on some parameters such as time, cost and technical facilities which vary from project to project.

## 3. Discussion and Conclusions

Since medical science is constantly evolving, the need to modify the database is permanent and the capacity to update the database without affecting the general model is important. Flexibility or extensibility is a major advantage for the data model. For instance the EAV model by adding new fields in the database doesn't require changes in schema. Adding new rows in a specific table can easily do it. But, the disadvantage of this model concerns mostly performance issues (e.g. time of performance, query complexity).

The data coverage of research questions together with data integration are the major parameters for model selection. Thus, the achieved results of the ontology-driven model [11-14] showed the adequacy of the semantic integration methods developed by the project consortium. One of the main benefits demonstrated by this approach is increased portability of the semantically formalized data sources. Though the ontological data integration was achieved, local semantic formalization was not fully interoperable. In the ontology-based integration system, the automatic mapping from global to local ontologies using first-order logic hinders logical consistency [15]. Consequently, various local ontologies were not completely resolved in the global model. However, regarding the flexibility criterion, in DebugIT, query templates must be defined centrally for each new data source.

The choice of the data model based on i2b2 [16-19] for the EHR4CR environment is rather rational [20]. The data model of i2b2 was easily adapted to EHR4CR needs since the patient-centric model matched the patient-centric scenarios of the project. The understandability made an i2b2 to be an admired tool for data integration and querying by clinical users. The semantic interoperability was achieved through external EHR4CR software such as terminology services. It is also worth to notice one of the differences between the DebugIT and EHR4CR environments is that in the latter one the query templates are defined locally. Moreover, the clinical data warehouse based on i2b2 data model relies on the relational database mechanism where the query construction to access data is less complex than for the ontology-driven model.

The data integration approaches such as ontology-driven and data-driven are able to homogenize the distinct data sources. In the EAV model due to the specificity of the project, the homogenization is done on demand.

Conclusively, the described projects have different goals and to define a suitable data integration model which would fit the needs of every project is not realistic.

# References

[1] Voss EA, Makadia R, Matcho A, Ma Q, Knoll C, Schuemie M, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. Journal of the American Medical Informatics Association. 2015:ocu023.

[2] Garde S, Knaup P, Hovenga EJ, Heard S. Towards Semantic Interoperability for Electronic Health Records--Domain Knowledge Governance for open EHR Archetypes. Methods of information in medicine. 2007;46(3):332-43.

[3] Lovis C, Colaert D, Stroetmann VN. DebugIT for patient safety-improving the treatment with antibiotics through multimedia data mining of heterogeneous clinical data. Studies in health technology and informatics. 2008;136:641.

[4] Schulz S, Karlsson D. Records and situations. Integrating contextual aspects in clinical ontologies. Bio-Ontologies 2011. 2011.

[5] De Moor G, Sundgren M, Kalra D, Schmidt A, Dugas M, Claerhout B, et al. Using electronic health records for clinical research: The case of the EHR4CR project. Journal of biomedical informatics. 2014.

[6] Teodoro DH, Choquet R, Schober D, Mels G, Pasche E, Ruch P, et al. Interoperability driven integration of biomedical data sources. Studies in health technology and informatics. 2011;169:185-9.

[7] Murphy S, Gainer V. Working Assumptions of i2b2 Data.

[8] Ouagne D, Hussain S, Sadou E, Jaulent M-C, Daniel C. The Electronic Healthcare Record for Clinical Research (EHR4CR) information model and terminology. Studies in health technology and informatics. 2012(180):534-8.

[9] Bache R, Taweel A, Miles S, Delaney B. An Eligibility Criteria Query Language for Heterogeneous Data Warehouses. Methods Inf Med. 2015;54(1):41-4.

[10] Moody DL, Shanks GG. Improving the quality of data models: empirical validation of a quality management framework. Information systems. 2003;28(6):619-50.

[11] Daniel C, Choquet R, Assele A, Enders F, Daumke P, Jaulent M-C, editors. Comparing the DebugIT dashboards to national surveillance systems. BMC Proceedings; 2011: BioMed Central Ltd.

[12] Schobera D, Choquetb R, Depraeterec K, Endersd F, Daumked P, Jaulentb M-C, et al. DebugIT: Ontology-mediated layered Data Integration for real-time Antibiotics Resistance Surveillance.

[13] Pasche E, Ruch P, Teodoro D, Huttner A, Harbarth S, Gobeill J, et al. Assisted knowledge discovery for the maintenance of clinical guidelines. PloS one. 2013;8(4):e62874.

[14] Choquet R, Daniel C, Grohs P, Douali N, Jaulent M-C, editors. Monitoring the emergence of antibiotic resistance using the technology ot the DebugIT platform in the HEGP context. BMC Proceedings; 2011: BioMed Central Ltd.

[15] Schulz S, Stenzhorn H, Boeker M, Smith B. Strengths and limitations of formal ontologies in the biomedical domain. Revista electronica de comunicacao, informacao & inovacao em saude: RECIIS. 2009;3(1):31.

[16] Boussadi A, Caruba T, Zapletal E, Sabatier B, Durieux P, Degoulet P. A clinical data warehouse-based process for refining medication orders alerts. Journal of the American Medical Informatics Association. 2012:amiajnl-2012-000850.

[17] Cuggia M, Garcelon N, Campillo-Gimenez B, Bernicot T, Laurent J-F, Garin E, et al., editors. Roogle: an information retrieval engine for clinical data warehouse. MIE; 2011.

[18] Mate S, Bürkle T, Köpcke F, Breil B, Wullich B, Dugas M, et al. Populating the i2b2 database with heterogeneous EMR data: a semantic network approach. Studies in health technology and informatics. 2010;169:502-6.

[19] Segagni D, Gabetta M, Tibollo V, Zambelli A, Priori SG, Bellazzi R, editors. ONCO-i2b2: improve patients selection through case-based information retrieval techniques. Data Integration in the Life Sciences; 2012: Springer.

[20] Vishnyakova D, Bottone S, Pasche E, Lovis C. Practical Implementation of a Bridge between Legacy EHR System and a Clinical Research Environment. Studies in health technology and informatics. 2014;197:29-33.

# Automated Data Aggregation for Time-Series Analysis: Study Case on Anaesthesia Data Warehouse

Antoine LAMER[a, b, 1], Mathieu JEANNE[a,b], Grégoire FICHEUR[c], and Romaric MARCILLY [b]

[a] *Univ. Lille, CHU Lille, Pôle d'anesthésie-réanimation, F-59000 Lille, France*
[b] *Univ. Lille, Inserm, CHU Lille, CIC 1403 - Centre d'investigation clinique Innovations Technologiques, F-59000 Lille, France*
[c] *Univ. Lille, CHU Lille, EA 2694 - Santé publique : épidémiologie et qualité des soins, F-59000 Lille, France*

**Abstract.** Data stored in operational databases are not reusable directly. Aggregation modules are necessary to facilitate secondary use. They decrease volume of data while increasing the number of available information. In this paper, we present four automated engines of aggregation, integrated into an anaesthesia data warehouse. Four instances of clinical questions illustrate the use of those engines for various improvements of quality of care: duration of procedure, drug administration, assessment of hypotension and its related treatment.

**Keywords.** Data Collection, Vital Signs, Anesthesia.

## Introduction

Operational databases daily collect high volumes of data in medical care centres and offer opportunities for reusing these data, e.g. for research purposes or in order to assess the quality of care [1]. For instance, analyses of anaesthesia time-series can be useful to identify the events that impact the patient outcome [2]. However, the structure of these databases is designed for clinical or administrative purposes, not for their secondary use [3]. Operational databases, in particular, collect high volumes of raw and heterogeneous data that cannot be reused directly in their initial format: for example, Anaesthesia Information Management Systems (AIMS) register data about the anaesthesia procedure mainly for medical follow up, and sometimes for legal purposes. Each anaesthetic procedure produces around two thousand measurements of vital signs (e.g. heart rate, arterial pressure) and one hundred events (e.g. surgical stages, administration of drugs). For clinical research purposes, it is important to assess specific events during predefined periods: for instance, the evolution of vital signs in relation to drug administration, the occurrence of some specific adverse events (e.g. hypotension, tachycardia) during anaesthesia or the cumulative dose of various drugs administered to the patient during the procedure. Despite the high volume of data collected by the AIMS, several quality

---

[1] Corresponding Author.

checks and various analyses must be carried out on these data before any clinical research can take place [4].

First, it is necessary to integrate data from heterogeneous sources into a common repository, such as a data warehouse: these well-known systems are used in order to manage high volume of data [4-5]. Then, after a cleaning and transformation process, data can be aggregated and synthesized in order to obtain relevant and meaningful information that can be used for clinical or statistical analyses. The aggregation process is based on business rules provided by expert clinicians.

The goal of this paper is to highlight opportunities related to data aggregation: turning raw data into substantial information for reuse. With this aim in mind, we describe four engines that have been developed to perform the automatic aggregation of anaesthesia time-series data and to carry out clinical or statistical analyses.

## 1. Methods

### 1.1. Available data

A data warehouse has been developed in the University Hospital of Lille to reuse data registered by the AIMS [6]. It stores data related to around 55 000 interventions performed each year (e.g. patient history, vital signs, various stages of the procedure, drug administration's) which then may be queried to fed aggregation engines.

### 1.2. Aggregation methods

The aggregation step consists in transforming data into meaningful information. *Data* are characterized by a high number of records (in rows) for a few numbers of attributes (in columns). On the contrary, *information* is characterized by a higher number of attributes (in columns) and a lower number of records. Therefore, the volume of data is decreased, while the volume of meaningful information is enhanced (Figure 1).
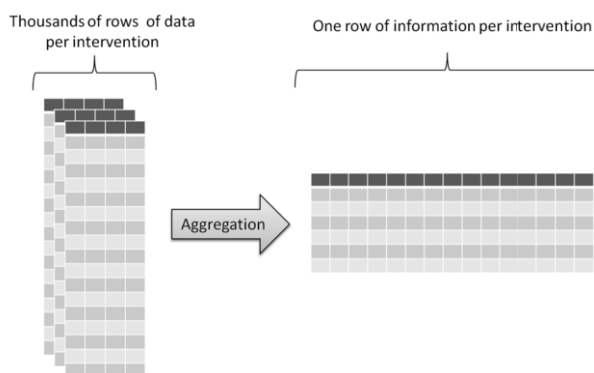


**Figure 1.** The aggregation process transforms raw data into aggregated data. Raw data are registered by source systems and present a high volume but a low information rate while aggregated data have a lower volume and more information.

We developed four aggregation engines to reuse anaesthesia time-series data (Figure 3). Each aggregation method was evaluated by experts by comparing results with source data. Engines can be used jointly according to the questions.

- The **"study periods"** engine is based on the detection and selection of predefined elements among the measurements, events, drugs (e.g. first non-null value, first administration of hypnotic drug) which correspond to the start or the end of specific periods (Figure 2, A1-A2) [7]. Example: anaesthesia or surgery stages, periods surrounding the administration of a given drug.
- The "**aggregated measures**" module employs aggregate functions (e.g. mean, median, minimum, maximum) to compute statistical indicators based on the measurements made during a certain study period previously defined (Figure 2, B1-B2). Example: mean of Heart Rate (HR) during anaesthesia, maximal value of mean arterial pressure within ten minutes after a vasopressor has been administered.
- The "**abnormal values of vital parameters**" module detects adverse events based on time-series of vital signs (Figure 2, C1-C2) [8]. These events are characterized by the total time elapsed outside a pre-defined range, the number of these episodes and the extreme (lowest or highest) value. Example: episodes of hypotension with Mean Arterial Pressure (MAP) lower than 60 mmHg.
- The "**drug administration**" module computes for each drug the total dose of the successive administrations within a study period (Figure 2, D1-D2). Example: total dose of propofol used during the induction period.
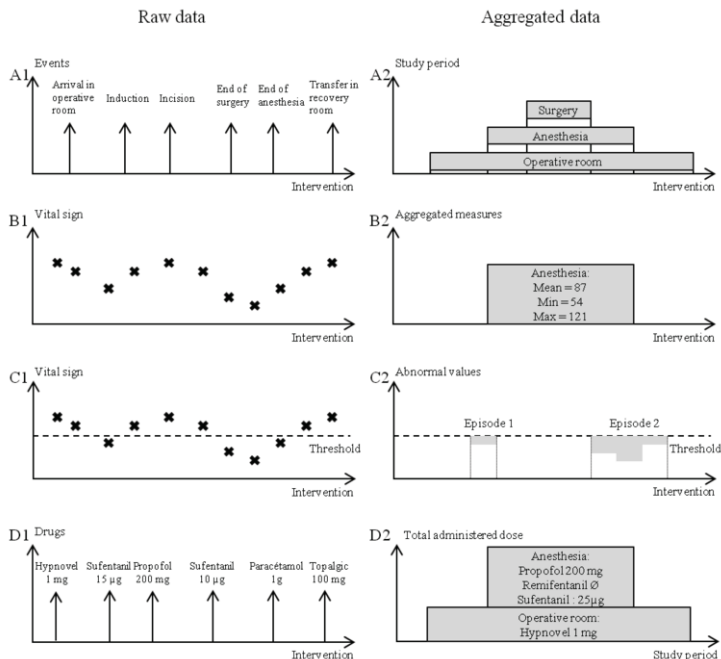


**Figure 2. A**ggregation engines. A) Study period ; B) Aggregated measures ; C) Adverse events ; D) Drug administration.

## 2. Results

In order to illustrate how the aggregated engines works, we have developed four study cases to know (i) what is the duration of the anaesthesia procedure, (ii) what are the variations of HR around administration of atropine, (iii) how can be characterized the drop of MAP after induction of anaesthesia with propofol and (iv) what is the total amount of ephedrine administered to manage blood pressure following the start of anaesthesia. From a clinical perspective, results for a given question can be regularly assessed over time to improve the quality of care A total of 276 775 anaesthetic procedures have been carried out at the University Hospital of Lille, France, between January 2010 and December 2014.

**(i) Length of anaesthesia procedure - Engine "Study period"**: the length of anaesthesia has been measured for 261 996 anaesthetic procedures leading to a mean (standard deviation) of 161 (200) min (cf [7] for more details).

**(ii) Evolution of heart rate - Engines "Aggregated measures" and "Study period"**: it has been calculated on 17118 interventions. HR evolves from a mean minimal value of 53 beats per minute (bpm) before administration of atropine, to a mean maximal value of 87 bpm after administration, and remains constant around 76 bpm after 15 minutes after the administration (figure 3).
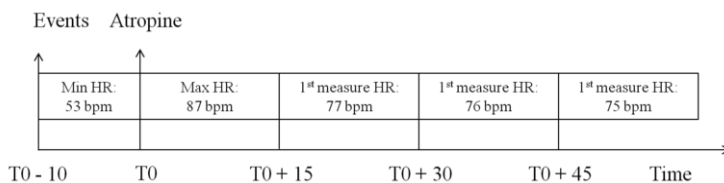


**Figure 3.** Aggregated measures of heart rate variations around the administration of atropine.

**(iii) Abnormal values of Mean Arterial Pressure (MAP) - Engines "Study period" and "Abnormal values":** it has been calculated on 81 014 interventions between induction and start of surgery. Table 1 presents, for various minimal thresholds of MAP, the number of interventions which dropped below the threshold, the mean duration of the first episode and the delay between the induction and the minimal value reached.

**Table 1.** Drop of Mean Arterial Pressure between induction of anaesthesia and start of surgery

| Minimal Threshold (mmHg) | Nb of interventions (%) | Median duration of first episode (min) | Mean (SD) time between induction and minimal value (min) |
|---|---|---|---|
| < 50 | 11846 (14.62%) | 3.05 | 16.91 (13.20) |
| < 55 | 10822 (13.36%) | 2.98 | 17.93 (13.84) |
| < 60 | 14016 (17.3%) | 3.67 | 18.17 (14.09) |
| < 65 | 14304 (17.66%) | 4.87 | 17.94 (14.09) |
| < 70 | 11807 (14.57%) | 4.95 | 17.79 (14.28) |
| < 75 | 8266 (10.2%) | 4.97 | 17.67 (14.33) |
| - | 9953 (12.29%) | - | - |

**(iv) Total dose of ephedrine - Engines "Study period", "Abnormal values" and "Drug administration"**: total dose of ephedrine, used as a treatment of hypotension

which may follow induction, has been calculated on 81 014 interventions. Table 2 presents, for each lowest thresholds of MAP, the number of interventions in which ephedrine has been administered between induction and start of surgery and the median (interquartile) of the dose administered.

**Table 2.** Administration of ephedrine between induction and incision

| Threshold (mmHg) | Interventions with administration of ephedrine following induction (%) | Ephedrine (mg) (median [interquartile]) |
|---|---|---|
| < 50 | 6566 (55.43%) | 9 [9 ; 15] |
| < 55 | 4443 (41.06%) | 9 [9 ; 12] |
| < 60 | 3747 (26.73%) | 9 [6 ; 9] |
| < 65 | 1850 (12.93%) | 9 [6 ; 9] |
| < 70 | 688 (5.83%) | 9 [6 ; 9] |
| < 75 | 261 (3.16%) | 9 [6 ; 9] |
| - | 177 (1.78%) | 9 [6 ; 12] |
| Total | 17732 (21.89%) | 9 [9 ;12] |

## 3. Discussion

In this paper, we highlighted how aggregation engines can be used to permit reuse of data produced by the AIMS. The aggregation of raw data increases information while decreasing the volume of data. Even if developing these engines is time consuming and requires informatics and clinical expertise and validation, once engines are developed, they can be quickly and easily customized to carry out new studies. Moreover, the four engines presented in this paper can be adapted to various time-series (e.g. data recorded on patients in intensive care) and can be joined when it is necessary to combine various types of information to answer a specific clinical question. Routinely performed in a data warehouse, automated data aggregation makes aggregates available, even for records recently integrated, and directly operable for analysis and retrospective clinical studies. These studies can be repeated over time for quality of care assessment and improvement.

## References

[1] M.G. Weiner, P.J. Embi. Toward Reuse of Clinical Data for Research and Quality Improvement: The End of the Beginning? Ann Intern Med. 2009 Sep 1;151(5):359–60.
[2] Sessler DI, Sigl JC, Kelley SD, Chamoun NG, Manberg PJ, Saager L, et al. Hospital stay and mortality are increased in patients having a "triple low" of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia. Anesthesiology. 2012 Jun;116(6):1195–203.
[3] K. Dentler, A. ten Teije, N. de Keizer, R. Cornet. Barriers to the reuse of routinely recorded clinical data: a field report. Stud Health Technol Inform. 2013;192:5–7.
[4] C.M. Nunez. Advanced techniques for anesthesia data analysis. Semin Anesth Perioper Med Pain. 2004 Jun;23(2):121–4.
[5] R. Kimball.The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing, and Deploying Data Warehouses. John Wiley & Sons (1998)
[6] A. Lamer, M. Jeanne, B. Vallet, G. Ditilyeu, F. Delaby, B. Tavernier, and R. Logier, "Development of an anesthesia data warehouse: Preliminary results," IRBM, vol. 34, no. 6, pp. 376–378, décembre 2013.
[7] A. Lamer, J. De Jonckheere, R. Marcilly, B. Tavernier, B. Vallet, M. Jeanne et al. A substitution method to improve completeness of events documentation in anesthesia records. J Clin Monit Comput. 2015.
[8] A. Lamer, M. Jeanne, R. Marcilly et al. Methodology to automatically detect abnormal values of vital parameters in anesthesia time-series: proposal for an adaptable algorithm. Comput Methods Programs Biomed. 2015. In press.

# Case-Based Learning: A Formal Approach to Generate Health Case Studies from Electronic Healthcare Records

Fabrizio L. RICCI[a,1], Fabrizio CONSORTI[b], Manuel GENTILE[c], Linda MESSINEO[c], Dario LA GUARDIA[c], Marco ARRIGO[c], and Mario ALLEGRA[c]

[a] *Institute for Systems Analysis and Computer Science and Institute of Molecular Bioimaging and Physiology, National Research Council of Italy, Italy*
[b] *Dept. Surgical Sciences, University Sapienza of Rome and SIPeM, Italy*
[c] *Institute for Educational Technologies, National Research Council of Italy, Italy*

**Abstract.** There is an increasing social pressure to train medical students with a level of competency sufficient to face clinical practice already at the end of their curriculum. The case-based learning (CBL) is an efficient teaching method to prepare students for clinical practice through the use of real or realistic clinical cases. In this regard, the Electronic Healthcare Record (EHR) could be a good source of real patient stories that can be transformed into educative cases. In this paper a formal approach to generate Health Case Studies from EHR is defined.

**Keywords.** Case-based learning, Electronic Healthcare Record, Experiential Learning

## Introduction

There is an increasing social pressure to train medical students with a level of competency sufficient to face clinical practice already at the end of their curriculum. Active methods of teaching/learning can guarantee the development of high quality expertise, because they are directed to the development of knowledge and also to its use in real or realistic settings. The case-based learning (CBL) is a teaching method that belongs to the more general class of inquiry or discovery learning [1]. The CBL has been defined as an approach to education, which aims to prepare students for clinical practice, through the use of real or realistic clinical cases. These cases link theory to practice through the application of theoretical knowledge to the cases themselves and encourage the use of methods of inquiry based learning [2]. The ideal features of a case for CBL are: a) be authentic (based on real patient stories); b) involve common scenarios; c) tell story, d) be aligned with defined learning outcomes; d) have educational value; e) stimulate interest; f) create empathy with the characters; f) include quotations in the patient voice to add drama and realism; g) promote decision making; and h) have general applicability.

---

[1] Corresponding Author.

Learning activities in CBL are essentially ways to "play" a simulation of management of a case, and it is very different from the traditional "case reports", in which clinical features and the decision-making process are illustrated to the "passive" learner.

The selection or production of cases for CBL is a time consuming activity, especially taking into account the time necessary to transform the extracted clinical information into a suitable format for the educational use. Direct cases generation from an electronic database were attempted in radiology, by connecting an educational system to the local PACS. The use of true medical records is not reported in the literature [3].

The Longitudinal Electronic Healthcare Record (L-EHR) could be a good source of real patient stories to generate cases for educational purposes. The L-EHR is a data collection of patient health information generated during all patient's encounters with caregivers in any care delivery setting. It integrates standard clinical documents produced by different specialists and by different organizations (e.g. laboratory and hospital information systems, GP's record); these documents have the clinical contact (during which it was written) and the health issue (the origin of the clinical contact) as metadata. The L-EHR can contain the clinical history of a citizen during his entire life.

Obtaining health case studies (HCS) from the L-EHR also helps to have, for each significant problem of health, a wide selection of presentations and clinical courses, probably similar but not identical. This is required for the development of a mature clinical expertise [4]. The extraction of HCS from the L-EHR for educational purposes must allow the learners to practice: a) to identify the essential information and the points of decision; b) to take autonomous diagnostic, therapeutic and management decisions; c) to compare their decisions with those actually made in the reality of that particular case; d) to balance the rationale of their decisions with the available scientific evidence, collected from knowledge bases; and e) to discuss their work with a tutor.

## 1. Longitudinal Electronic Healthcare Record as an information source for CBL

Each episode of care starts from a clinical problem (e.g. symptom) and then is described by the evolution of this problem along with the clinical events carried out during the care. According to CONTsys "the health issue is the reason for the request for health care"[5]. The health issue is a label assigned by a health professional who participates in the care process in order to describe the patient's health condition, at a given time. The label is subjective because it always depends on the observer point of view (e.g. other professionals could use another label); moreover, over time, the same observer could assign different label for the same condition.

The clinical history of a patient can be described as a sequence of contacts, and also as a network of health issues that provides the complete semantics of the various problems encountered by the patient and their evolution in time. The edges of this graph have different meaning, for example: a) in depth (the problem is analysed in depth and defined more); b) complication (the problem degenerates into a different issue); and c) consequence (some diseases, especially chronic, could produce complications).

In order to classify a health issue one of the most popular classification codes could be used: ICPC (International Classification of Primary Care), SNOMED (Systematized Nomenclature of Medicine) or ICD (International Classification of Diseases).

## 2. The generation of a Health Case Study.

The extraction of the HCS is realized by selecting the entire path of treatment starting from the initial healt issue. It includes the events that compose the path of care as well as all of the problems associated with it, independently of the doctor who defined them. The construction process of the HCS occurs through the following phases:

➢ Extraction and anonymisation: a) the selection of the initial health issue from which all the care pathway must be derived; b) export of selected case; and c) anonymization in according to national laws [5].

➢ Preparation of the HCS: a) the inclusion of comments and notes; b) the connection of the elements of the pathway with external documents, websites and other resources; and c) the transformation of the HCS( e.g., to extract, namely the generation of a sub-HCS; to generalize, i.e. to change a detailed information to its type; to change the absolute times to the relative times).

Obviously, the HCS generated has to be contextualized in the light of a concrete learning path in orde to become a learning resource.

## 3. The formalization of the health case study

The formalization of the HCS is based on the formal representation of the L-EHR. The HCS is a complex artefact composed by: a) network of health issues, formally a connected graph; b) the clinical events linked to the network of the health issues; and c) the clinical documents linked to the network of the health issues.

Let: a) $P$ a set of Health Issues; b) $ass: P \rightarrow P$ a function that link a Health Issue to another Health Issue; then the network of Health Issues is a pair $\varphi$: $\varphi = < P, ass >$.

Moreover, let: a E a set of clinical events; b) D a set of clinical documents; c) $e_p: P \rightarrow 2^E$ a function that return the clinical events related to the episode of care linked to a specific health issue; d) $g_n: P \rightarrow 2^D$ a function that return the clinical documents produced during all the clinical events linked to a specific health issue by means of the $e_p$ function; then L-EHR ψ is a quintuple $\psi = \langle \varphi, E, D, e_p, g_n \rangle$.

Let F the set of possible L-EHR.

The extraction of the health case study is based on the choice of the initial health issue and then on the extraction of all the all the directly and indirectly connected health issues together with the related clinical events and clinical documents.

The first step is the extraction of the related health issues. Given a health issues network $\varphi = < P, ass >$, the function $path: P \rightarrow 2^P$ is a function that return all the health issue directly and indirectly connected with the initial health issue.

Let $p \in P$ as the initial health issue,

$path(p) = \{p\} \cup \{x\ P: (\exists y \in path(p): y = ass(x) \lor x = ass(y))\}$ ⊆P

The network of health issues of the HCS related to the initial problem p is

$\varphi_p = < path(p), ass|z \in path(p) >$.

The extraction of the HCS from the initial health issue $p$ is based on the following function $ext(p): P \rightarrow 2^F$.

Therefore, given a L-EHR $\psi = \langle \varphi, E, D, e_p, g_n \rangle \in$ F and a Health Issue p $\in$ P, the HCS $\psi_p$ is $\psi_p = ext(p) = \langle \varphi_p, E_p, D_p, e_p | E_p, g_n | D_p \rangle \subseteq 2^F$; where: a)

$E_p = \bigcup_{y \in path(p)} e_p(y)$ ⊆ E ; and b) $D_p = \bigcup_{y \in path(p)} g_n(y)$ ⊆ D.

## 4. The use of the health case study

Kolb [7] defines learning as "the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience".

The case based learning allows trainees to activate experiential learning in a more effective way if the cases are based on a real scenario. According to this approach, the analysis of a HCS generated from L-EHR could be a useful tool for building a learning environment; this approach can be applied in the field of training of general practitioner as well as for specialists to allow them to address the complexity of problems that arise in their daily activities.

There are at least three different levels of details that offer a different vision of the answer, depending on which aspects of the learner wants to deepen. These levels, given in order of extent and depth of the response, are: a) the diagnostic process in which the recognized guidelines can be directly used; the goal of the exercise is to train the use of guidelines b) the clinical reasoning, mainly hypothetical-deductive reasoning; the goal is to train clinical argument skill to support the path that goes from identification to interpretation of information and therefore differential diagnosis; and c) deepening in terms of pathophysiological mechanisms, where the goal is to teach in a more systematic way epidemiology, prognostic, therapeutic (medical and surgical) and diagnostics.

The additional external information are therefore characteristics of each level, such as for example: a) database consisting of guidelines (e.g. Cochrane library); b) individual studies, clinical trial; and c) biochemistry at the base of the physio-pathological aspects.

Regarding the metadata describing the HCS must also take account of the presence of educational metadata: a) competence (level 1 of Tuning)[8]; b) work activity (level 2 of Tuning)[8]; and c) theoretical knowledge / factual or use specific skills brought into action [9].

As an example, an educational goal like "how to make a diagnosis," the description of an HSC result in the presence of the following metadata:

- competence: evaluate a clinical presentation and make the differential diagnosis;
- work activity: to recognize the elements of a clinical image, prescribe diagnostic tests, interpret and make the differential diagnosis;
- knowledge: epidemiology and pathophysiology, clinical - it is placed at lower levels of Bloom's taxonomy: remember the terms, facts, mechanisms, criteria;
- skills: correlation between elements of knowledge - it is placed at high levels of Bloom's taxonomy: analyse, interpret, reorganize.

## 5. Conclusions

In this paper propose a methodology to support the creation of HCS from a L-EHR. The basic idea and its implication from an educational point of view were discussed with experts from the Italian Society for Medical Education (SIPeM). The main value of generating HCS from L-EHR is in the availability of a very wide set of complex cases, representing the real-world situation. According to Kolb [7] the activation of learning through experiences requires to face increasingly difficult challenges. If Ad hoc, manually generated cases can be a starting point and bedside teaching with a direct

involvement of student the final goal, the management of complex, real-life cases, automatically generated from an a L-EHR is a valuable intermediate objective [10]. In order to support this process a formal description of the HSC was defined and some use cases were proposed.

The implementation of the proposed approach requires to use an L-EHR system in which both, the conceptual model and the implemented features allow the implementation of the designed tools.

A prototype of the proposed system has been developed in the framework of the project Smart Health 2.0 funded by the Italian government research and innovation programme "Smart Cities and Communities and Social Innovation". The prototype allows user to skim through then information contained in the L-HER to collect the clinical events considered essential to the definition of the base case. These elements can then be re-processed in a causal network according to the model proposed in this paper to finalize the HCS. The developed features are being tested since June 2015 and the first results will be available before the June 2016.

## References

[1] Banchi H, Bell R. The many levels of inquiry-based learning. Science and Children, 2008; 46:26-29

[2] Thistlethwaite JE, Davies D, Ekeocha S et al.. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. Med Teach. 2012;34(6):e421-44

[3] Dugas M, Trumm C, Stäbler A, Pander E, Hundt W, Scheidler J, Brüning R, Helmberger T, Waggershauser T, Matzko M, Reiser M. Case-oriented computer-based-training in radiology: concept, implementation and evaluation. BMC Med Educ. 2001;1:5.

[4] Norman G, Bordage G, Page G, Keane D. How specific is case specificity? Med Educ. 2006 Jul;40(7):618-23

[5] "System of Concepts to Support Continuity of Care (CONTsys)." [Online]. Available: http://www.contsys.net/. [Accessed: 21-Dec-2014].

[6] Schema di decreto del presidente del Consiglio dei Ministri: fascicolo sanitario elettronico, ai sensi dell'art. 12, comma 7 del decreto-legge 18 ottobre 2012 n. 179, e dell'art. 13 comma 2 quater del decreto- legge 21 giugno 2013 n. 69 (febbraio 2014)

[7] Kolb, D.A. (1984): Experiential learning: experience as the source of learning and development. Englewood Cliffs, NJ: Prentice Hall.

[8] Cumming, A., & Ross, M. (2007). The Tuning Project for Medicine--learning outcomes for undergraduate medical education in Europe. Medical Teacher, 29, 636–641. doi:10.1080/01421590701721721

This page intentionally left blank

# Posters

This page intentionally left blank

# Retrieving Clinical and Omic Data from Electronic Health Records

Chloé CABOT [a,1], Romain LELONG [a], Julien GROSJEAN [a], Lina F. SOUALMIA [a,b], and Stéfan J. DARMONI [a,b]

[a] *SIBM, Rouen University Hospital & TIBS, LITIS EA 4108, Rouen, France*
[b] *INSERM, LIMICS UMR 1142, Paris, France*

Integrating omic and clinical data within Electronic Health Records (EHR) can lead to significant progress in the development of new diagnostic tests and therapies as well as improve our understanding of complex genetic diseases and cancers. Several tools and frameworks have been proposed for querying and using EHRs data such as Integrating Biology and the Bedside (i2b2), STRIDE and EMERSE. Here we propose an information retrieval system which allows querying both molecular biology (*omic*) and clinical data in EHRs.

Clinical data has been obtained from Rouen University Hospital. Nearly 2,000 EHRs and 200,000 medical records have been integrated within our clinical and omic data model. Omic experimental data has been obtained from local collaborations and The Genome Cancer Atlas. Terminologies and ontologies are provided by the Health Terminology / Ontology Portal (www.hetop.eu).

The provided system allows to query both clinical and omic data in one or several EHRs. Queries can be processed at a patient level, a stay level or an experimental study level. Multiple omic data types can be retrieved: expression data (gene, protein, miRNA), genomic variants, copy number variants, comparative genomic hybridization and DNA methylation data. Terminologies and ontologies such as SNOMED, MeSH, ICD10 are used as references within the graphical user interface and can be used to build queries. For example, all patients having genomic variants on genes annotated with the Gene Ontology term GO:0042246 "Tissue regeneration" can be retrieved with the query

```
patient(variant(gene(GOterm(label="tissue regeneration")))).
```

While the reference solution i2b2 is widely adopted in both academic community and industries, our system brings some decisive advantages. In fact, it can handle more use cases mixing symbolic, textual, numerical and chronological data and many clinical and omic data types. Moreover, it is highly extensible and adaptable to future omic data. Furthermore, terminologies and ontologies integrated within this system can provide a useful support to information retrieval. An ongoing project is to evaluate in the coming weeks the usability of the search tool by a group of physicians and researchers.

---

[1] *Corresponding Author.*

# Smart Medications & the Internet of Things

José COSTA TEIXEIRA [a,1] and Catherine CHRONAKI [a]
[a] *HL7 Foundation, Brussels, Belgium*

In our digital society, where "smart things" become integrated and personalized, digital enhancements can increase safety of medicinal products and help avoid harm for patients. Accurate identification of medicinal products throughout their lifecycle, into prescription, dispensation, and compliance is essential. Wrong use of medication, or use of the wrong medication, can be very harmful. Integrating health information technology with IoT enables among others, individualized drug information, indications of drug-to-drug interaction, allergy alerts, tracking of medication use and effects in wellness and disease, and reporting of adverse events. Areas where standards support effective use of medicines are explored. Standards are key to integration of the IoT with eHealth. Creating an ecosystem of 'smart' connected health apps and sensors to enhance the use of medication requires standards for semantic and functional interoperability, respecting security and privacy. Healthcare standards support safety and quality in the interaction between the physical medication and the digital world. In one example, standards (GS1 for barcodes and HL7 for data capture and exchange) support medication identification across borders, and accessing drug databases to deliver personalized patient instructions and alerts. From the prescription, a pharmacist dispenses the right medication using barcodes and medical history of the patient. Accessing an HL7 FHIR ReSTful API ([www.HL7.org/FHIR](www.HL7.org/FHIR)), the pharmacy obtains the relevant product characteristics, considering the personal context of the patient: contraindications, allergies, interactions with other treatment, wellness data and habits. A "smart" pill box can help patients and their doctors monitor adherence to the prescribed schedule. openMedicine ([www.open-medicine.eu](www.open-medicine.eu)) provides a wallet or portfolio of identifiers to enable safe identification of medication throughout their lifecycle. Another case is chronic disease (diabetes) management, where patient monitoring is enhanced using integration standards between several 'smart' objects - weight scales, activity monitors, insulin pumps, and medication. Standards-based integration enables a 360º view of the patient treatment, "feeding" data from the digital world of the IoT to enable responsive treatment. Such dense integration is facilitated by the use of standards and exposes need for secure trustworthy device and data sharing.

Proper identification of medication across applications and borders is necessary to ensure safety. The IoT provides a lot of data, but this information must be properly captured and managed safely, so that it does no harm and privacy rights are observed. Working at the integration of eHealth and IoT, standards should continue to provide guidance for connecting "smart" things and medication helping people navigate the complexity of interconnected health systems with safety and trust.

---

[1] Corresponding Author.

# Interactive Videos within an e-Learning Context

*Jean-Charles DUFOUR[a,b,c,1] , Alain EMERIC[a] and Roch GIORGI[a,b,c]*
*[a]Aix-Marseille Université, UMR_S 912 (SESSTIM), IRD, 13385 Marseille, France*
*[b]INSERM, UMR_S 912 (SESSTIM), 13385 Marseille, France*
*[c]APHM, BIOSTIC, Pôle de Santé Publique, Marseille, France*

The "*Mediterranean-South e-Master of Public Health*" project aims at providing a high quality education program in Public Health, covering three complementary specialties: "Expertise and Engineering of Health Information Systems (EISIS)", "Quantitative and Econometric Methods for Health Research (MQERS)", "Public Health, Society, Development (SPSD)". In this respect, we have developed an extended and enriched pedagogical model using Information and Communication Technologies in Education (ICTE) including distance-learning programs with synchronous courses and asynchronous activities. The use of video is known to be adapted for e-learning. However we hypothesize that changes need to be introduced in this concept in order to enhance its pedagogical capacities and also to be in line with the new profile of today students, who are strong daily video consumers. Our objective is to facilitate the production of interactive videos meeting teaching objectives and which can be adapted to the different levels of students' knowledge. We have therefore elaborated three interactive video "models": Model 1: basic video enhanced with limited interactivity (chaptering to ease the navigation, associate content downloading…); Model 2: video or screencast with advanced interactive elements (quizzes, click-areas, annexed videos, access to key points in the video timeline depending on the student's actions…); Model 3: scripted videos with adaptive content (possibility to introduce additional animations, interactive elements conditioning the reading of video sequences…) according to the knowledge and the skills of the student. The formalisation of such videos can be complex. Therefore, we have developed a software to help the teacher (accompanied by a pedagogical engineer), with the elaboration and the graphical representation of an e-learning scenario.

The software produces a reusable file for the script of interactive videos. Today students are "digital and video" native. They make an extensive use of video in their daily live. However, educational videos cannot carry out their pedagogical purpose when only offered in a passive way. An active participation of the student is important in order to increase his/her attention, knowledge and skills. The use of interactive videos, adapted to the level of the student, is a new and innovative concept which meets the expectations of both teacher and student. Our work is a step towards to facilitate the production of such videos. It provides an e-learning tool and models allowing for the production of personalized learning media.

---

[1] Corresponding Author.

# Internet of Things Based Medication Adherence Assessment

Hubert EBNER[a,b,1], Dieter HAYN[a] , Martin KROPF[a] , Robert MODRE-OSPRIAN[a] and
Guenter SCHREIER[a,b]

[a]*AIT Austrian Institute of Technology, Digital Safety & Security Dept, Graz, Austria*
[b]*Graz University of Technology, Institute of Knowledge Discovery, Graz, Austria*

Non-adherence to drug therapy may seriously impact therapeutic goals. Consequences to patients are numerous. Today, the availability of mobile health (mHealth) systems allows collecting data directly from patients, anytime and anywhere [1]. Nevertheless, such mHealth systems do not yet consider confounding factors that influence medication adherence.

We describe a method for monitoring medication adherence using smartphones, a mobile app, and Near Field Communication (NFC). The system was developed to capture data from various sources and monitor patients' adherence based on an Internet of Things (IoT) architecture [2]. Using smartphones, it was possible to collect accompanying data directly from the patients' personal environment – "anywhere" and "anytime". These data were collected from various domains: health (medical devices), adherence (smart blisters), environment (weather), and lifestyle (physical activity and geo-fencing), i.e. "anything" – using on-board sensors (e.g. NFC, GPS), and external services (e.g. publicly available cloud data, Google services).

A smartphone with the newly developed mobile app was used as a user-centred terminal. The app provided the link between the targeted technologies, various data sources, and the user in the IoT-based mHealth system. The system was successfully tested in a small-scale feasibility study, showing that the app helps monitoring medication adherence and possible confounders in an intuitive way.

The developed IoT-based medication adherence assessment system does not only track medication intake, it also monitors several confounders, which are likely to influence the adherence pattern. Future work will focus on which data sources are most valuable for monitoring an individual patient's medication adherence.

[1] S. Stegemann, J.-P. Baeyens, F. Cerreta, E. Chanie, A. Lofgren, M. Maio*, et al.*, "Adherence measurement systems and technology for medications in older patient populations," *European Geriatric Medicine,* vol. 3, pp. 254-260, AUG 2012.
[2] G. Schreier, "The internet of things for personalized health," *Stud Health Technol Inform,* vol. 200, pp. 22-31, 2014.

---

[1] Corresponding Author.

# Recruiting Participants to Local Clinical Trials using Ontology and the IoT

Peter L. ELKIN[a,1], Daniel R. SCHLEGEL[a] and Edwin ANAND[a]
[a]*Department of Biomedical Informatics, University at Buffalo, Buffalo, NY USA*

Clinical trial recruitment is challenging. The pressure of clinical practice is such that often clinicians do not have time during their appointments to look up clinical trials and explain them to their patients. Meanwhile, patients have a strong interest in knowing what trials are available in their local area. Although national resources exist such as clinicaltrials.gov, participants do not find it easy to find studies in their local area or to find out how far they would have to go to meet a study coordinator or for their study appointments. Community engagement has been suggested as one solution to this problem. We are designing smartphone applications that allow physicians and patients to easily evaluate studies, the study schedule, and the location of the site closest to their home to determine their level of interest. Searching for studies can be done by typing in their disorder or choosing the type of trial they are interested in. When patients find a study which interests them they just push a button and their contact information is sent to the study coordinator who can contact them to begin the recruitment. Self-recruitment has the potential to speed enrolment and help us translate basic research into new therapies for our patients.

Clinical trials in our region will be registered into one of two Clinical Trial Management Systems. These systems will, on a regular basis, send registered clinical trials to a database backing web services. Our applications will access the clinical trial data via these web services. The data associated with a trial, such as the title, abstract, inclusion and exclusion criteria, will be codified with ontology terms using our Natural Language Processing (NLP) technology for easy search by prospective patients.

As we are still building these tools, we have not yet evaluated them. To evaluate the success of such a tool, it is necessary to measure the satisfaction of participants, study coordinators, and principal investigators, as well as the rates and geographic distribution of recruitment. It is important that the system be useful, but not intrusive or annoying.

Nationally and internationally, we currently have a serious problem with clinical trial recruitment. Information such as geographic data and study calendars, which can be downloaded from our clinical trial management system, decrease practical uncertainty for potential recruits. Speeding the time it takes to recruit patients to clinical trials has the potential to improve our rate of recruitment to clinical trials.

---

[1] Corresponding Author.

# Design and Evaluation of a Multi User Medication Reminder Mobile App

MinaFALLAH[a,1] and Mobin YASINI[b]

[a]*School of Allied Medical Science, Tehran University of Medical Sciences,Tehran, Iran*
[b]*DMD santé, Research and development department, Paris, France*

Medication reminder systems could improve medication adherence particularly in older adults and enhance their disease management. This would be more serious when the medicines are vital.

We designed a mobile medical app based on Android platform: firstly, the use case scenarios have been conceptualized in collaboration with a health professional and an older adult used to take medications daily. Unified Modelling Language was used to model the use cases. Once the model had been developed, a medical doctor who did not participate in the development of the model validated the model obtained.

The evaluation will be performed in two parts: 1) Usability evaluation using System Usability Scale (SUS) method. 2) Effectiveness evaluation by designing a survey asking the target users to score the impact of the application on an analogical scale questionnaire. The users will be asked to complete the questionnaires after handling the app for at least 10 days.

Health professionals successfully tested the app designed for its effectiveness. Examples of the app functions are notification to the user to take his/her medication (every 15 minutes until 2 hours), managing the drug stock at home, calculation of the number of medications to take for a short trip and sending reminder messages to the entourage of the patient. Therefore, the family and friends can take part to this adherence. The questions included in the efficacy evaluation are listed briefly below:

a) Compared to the period when I did not use the app, the number of missing medication seems to be reduced (0 for no, not at all and 10 for yes, definitely).

b) Compared to the period when I did not use the app, I miss less my medications but I take them sometimes with a delay which is generally more than 1 hour (0 for totally disagree, 10 for totally agree)

Usability and performance evaluations are in progress and seem promising.
Respecting the ergonomic aspects as well as the involvement of the health professionals and target users in the conception and development of the health related apps, would lead to create tools that are more likely to succeed with good results once they are subject to a rigorous evaluation.

---

[1] Corresponding Author.

# Barriers to the Success of Health ICT Implementations – A Report from Norway

Conceição GRANJA[a,1], Gro-Hilde ULRIKSEN[a], Kari BYD[a], Stein Roald BOLLE[a] and Gunnar HARTVIGSEN[a,b]

[a] *Norwegian Centre for eHealth Research, University Hospital of North Norway, Tromsø, Norway*

[b] *Dept. of Computer Science, UIT – The Arctic University of Norway, Tromsø, Norway*

**Keywords.** Hospitals, Health Personnel, Technology, Local Government, Electronic Health Records.

Healthcare providers are pressured to become more efficient and cost-effective. The promise of improved quality, greater access to care, timeliness of service and costs savings in healthcare provision has led to swift development of health Information and Communication Technology (ICT) services worldwide. In 2009, the Norwegian Ministry of Health and Care Services has initiated a reform of the health care sector through the Coordination Reform. In the Coordination Reform, increased collaboration with patients as active participants through ICT solutions is defined as a priority area. In an attempt to achieve a timely solution, health authorities and vendors are making enormous investments in health ICT. DIPS ASA, the largest Scandinavian health ICT vendor, is developing a new Electronic Health Record (EHR), DIPS Arena, where the aim is to fulfil the Norwegian Government demands on health and care services provision.

Despite the recognition of the enormous potential to improve the sector, health ICT has been failing to demonstrate its foreseen benefits, being its involvement in care processes limited to specific fields. This work explores the impact that ICT implementations have on health care organizations, with primary focus on the barriers to health ICT implementation that have been recognized to have origin in technology.

A MEDLINE review focusing on unsuccessful implementation due to technology issues was conducted. The main barriers were: unclear definition of the advantages of technology for the user (e.g. clinicians, nurses), lack of usability, and integration of data generated by health ICT devices into the EHR.

Interoperability is a main topic when designing health ICT. There is no use to collect health data if it is local and not available to health personnel. The use of data standards ensures that health data can be integrated in the EHR and shared among health entities.

The participation of the end-user in the design process provides useful insights on functionality and integration in daily practice, and contributes to the solution usability.

An important factor in the success of health ICT is the knowledge that the user has of the system. User training before implementation, and support during the systems' production, should not be neglected. Training and support should be considered as fundamental matter when preparing the technology for full-scale deployment.

---

[1] Corresponding Author.

# Cloud Computing in Healthcare:
# A Space of Opportunities and Challenges

Anežka HUCÍKOVÁ[a1] and Ankica BABIC[b]

[a] *Dep. of Information Science and Media Studies, University of Bergen, Norway*
[b] *Dep. of Biomedical Engineering, Linköping University, Sweden*

Cloud computing is a model of self-service on-demand network access enabling delivering computing resources and services. Contemporary research shows that around 50% of healthcare companies from large hospitals to ambulatory services across the US and Europe have already deployed cloud technology. Since the technology is simultaneously evolving, more healthcare companies are expected to move their enterprise communication to the cloud. There is a need for a more complex understanding of opportunities and challenges for technology providers and healthcare organizations. Secondary, the poster aims to suggest how such specific challenges can be dealt with.

A systematic literature review conducted in October 2015 considered 33 selected full text articles published from 2010 to 2015 and available in the MEDLINE database. The focus was to properly identify the pros and cons of cloud migration with the focus on healthcare. The results were summarized using qualitative analysis.

For both opportunities and challenges, several common categories such as management, security, technology, legal were identified. Further findings concerned the areas where the research on cloud computing has already undergone significant development (bioinformatics, genomics) or is expected to grow in the future (nursing).

In terms of opportunities, cloud technology offers, in comparison to conventional computing, more effective cost of new IT infrastructure and its maintenance, scalability and flexibility. A cloud solution is believed to reduce electronic health records (EHR) expenses, networking software and licensing fees and therefore encourages cloud deployment. Another opportunity can be seen in sharing big data within healthcare consortia or using cloud as a platform for e-Health applications.

Major challenges come with security and privacy concerns, vendors' compliance, data ownership, network latency and accessibility. Even though security concerns are expected to decline, deploying cloud still causes organizational concerns about rights management or data transfers. Potential users need to be instructed and trained on how to use cloud services to achieve maximum benefits and to work more effectively and flexible. Specifically designed learning and training tools (e-learning, customer immersion experience, workshops) can be used to overcome such personal obstacles.

---

[1] Corresponding Author.

# Development of a Web Portal Using Open Source Information Visualization Libraries

Georgy KOPANITSA[a,b,1]

[a] *Institute Cybernetic Center, Tomsk Polytechnic University, Tomsk, Russia*
[b] *Tomsk State University for Architecture and Building, Tomsk, Russia*

**Keywords.** Open source, web portal, diabetes, usability

An efficient user interface plays a major role in the implementation of personal health records. Some patients, especially elderly, require for instance a simple tool to control their health status. However, user interfaces are usually developed according to a single implementation, and it takes time and money to adapt it when the EHR dataset is changed. To develop interfaces to answer user requirements, we have conducted a survey with doctor interviews and patient focus groups from September 2013 to July 2014. Focus group participants were eligible if they were current home health consumers (patients older than 61 year-old that received home care within the previous three years). After the analysis of focus groups and the interpretation of interviews, we have developed a web portal based on d3.js open source visualization solutions.

In a second step, we conducted usability testing interviews (from January to February 2015). For this phase, participants were eligible based solely on age (≥65 year-old), regardless of previous experience with home care. During the usability testing interviews, the facilitator asked participants to test the electronic prototype and to describe their thoughts and reactions aloud, using a cognitive-based testing approach. After the portal has been implemented we have conducted usability testing interviews with 14 patients. The patients demonstrated confusion when they tried to use the portal. Many sat in front of the computer and removed their hands from the keyboard or mouse, either silently, demonstrating their unease, or explicitly asking how to proceed through the application. The facilitator prompted consumers to navigate through the portal, advising them how to accomplish each step while eliciting their thoughts. To improve the web-portal and make it more suitable for users, we could modify its sequence, navigation, and function during usability testing. Thus the portal underwent iterative modifications to respond to user suggestions and increase ease of use. In the early iteration of the design modification, usability testing demonstrated that consumers were confused and unsure about how to use the application, often stopping and looking for the facilitator to help. Many were hesitant to experiment by clicking on options or navigating between web pages, preferring to wait for instructions or to take an action only when they were certain of the outcome. The application of open source libraries to standardized medical data facilitates the development of user interfaces due to, on the one hand, the standard functions that are provided by libraries, and, on the other hand, the standard structure of data.

---

[1] Corresponding Author.

# Midwifery Education Introduce the Internet of Things

Martina MAVRINAC[a,1] , Gordana BRUMINI[a,b] and Mladen PETROVEČKI [a,c]

[a] *Department of Medical Informatics, Faculty of Medicine, University of Rijeka, Croatia*

[b] *Department of Basic Medical Sciences, Faculty of Health Studies, University of Rijeka, Croatia*

[c] *Department of Clinical Laboratory Diagnostic, Dubrava Clinical Hospital, Zagreb, Croatia*

**Keywords**. Education, Internet, Medical Informatics, Midwifery.

The Faculty of Health Studies, University of Rijeka in Croatia introduced a new course called *Evidence based midwifery (EBM)* for the university study course of Midwifery in 2015.

The learning objectives of the course are to improve medical education using evidence based medical databases researches and to introduce the emerging topic of the Internet of Things (IoT). The goal is to follow the improvement and frequency of use of e-health tools by students during the next five years.

At the basic level the course instructors introduced students to the definitions, origins and purposes of evidence-based medicine, evidence-based midwifery and highlighting e-health applications as up-and-coming tools in midwifery.

Through the course various examples of IoT in midwifery are provided to explain how e-health tools works providing smarter and more efficient experiences to users (patients).

At the beginning of the course no one of the 25 students were introduced to EBM and no one student know what IoT means (N=25) and have never used an e-health tool.

All midwifery students (N=25) were for the first time introduced in using e-health tools for patients: surveillance of expectant mothers (1), controlling medicines (2), baby monitoring (3) and reminder bottles (4).

Students used individually a tool for surveillance based on monitoring conditions of expectant mothers at home and in hospitals on their Smartphone (N=17).

All students learned how to use e-health tools improving patients' surveillance.

All students who attended the course expressed their satisfaction with the new gained knowledge in the field of medical informatics.

The integration of real world application of IoT into midwifery education is a very important topic improving students' knowledge and ICT skills.

---

[1] Corresponding Author.

# MedBioinformatics: Developing Integrative Bioinformatics Applications for Personalized Medicine

Miguel A. MAYER[a,1], Laura I. FURLONG[a], Ferran SANZ[a]
and the MedBioinformatics Consortium

[a]*Research Programme on Biomedical Informatics (GRIB), IMIM-Universitat Pompeu Fabra, Barcelona, Spain*

Progress in healthcare and biomedical research involves taking advantage of the huge amount of clinical data and biological knowledge that already exists and is currently being generated. Bioinformatic methods and tools have become essential for data extraction, management, and integration. However, it is necessary to overcome two barriers: the deficit of integrative computational approaches that combine different types of data and sources, and the lack of active involvement of the potential users in the process of creating applications for the management of biomedical information.

The general objective of the MedBioinformatics project is to develop integrative bioinformatics tools and software applications to be used by translational scientists and clinical practitioners. Oncology and Central Nervous System (CNS) disorders are the case studies selected based on their impact in terms of citizen wellbeing, the diversity of comorbidities, and their molecular basis being currently an active area of research. Algorithms linking genotypes and phenotypes, the integration of databases on known genetic interactions, and biomarker data will be integrated in a desktop-able application with a steep clinician learning curve. A Disease Trajectory Comorbidity Browser will provide quantitative risk measures for comorbidities based on disease prevalence data, their related molecular features, and phenotypic information.

The expected impacts are to accelerate the translation of results to clinical research, the widespread dissemination of new bioinformatics methods to maximise the accessibility and utility of biomedical data in research and medicine, and to increase innovation opportunities of commercial products in bioinformatics research in Small & Medium Sized (SME) enterprise-intensive fields in Europe.

---

[1] Corresponding Author.

# Incorporating Pharmacy Dispensing Records into Medical Records: Usability Challenges

Aude MOTULSKY[a,1] Isabelle COUTURE[b], Daniala WEIR[a] and Robyn TAMBLYN[a]

[a] *Clinical and Health Informatics Research Group*, Montreal, QC CANADA
[b] *McGill University Health Center;* Montreal, QC, CANADA

Incorporating data generated by the dispensation of medications in pharmacies into clinical processes of prescription and medication reviews at different points-of-care has revealed unexpected challenges. In 2013, Quebec implemented an application (SQIM) to allow clinicians to access current and past medications for all patients in the province. This study evaluated the accuracy and usability of this application for documenting the list of current medications for patients at admission to hospital

An observational study was conducted from June 2014 to January 2015 at a tertiary care hospital in Montreal, Canada. Patients taking more than 3 medications, who were over 65 years' old who arrived in the emergency department during the study period were eligible. The list of current medications from the SQIM was compared to the list obtained from community pharmacies by fax. The status of each medication (current or not) was compiled based on the date of the last dispensation. Any discrepancy between the two lists was documented, along with reasons for discrepancies. Descriptive characteristics were estimated.

One hundred and eleven (111) patients were included (mean age 76; 51% female), with a mean number of 11 (SD 6.3) active medications upon admission. A total number of 442 discrepancies were observed in 71 patients (mean number per patient 6.5, SD 6.0). Three types of discrepancy were observed: a) 44.6% occurred when a medication was listed as current in the SQIM but should not have been (false positive); b) 43.9% when a medication was not listed as current in the SQIM but should have been (false negative); c) 11.5% were duplicate medications in the SQIM. 67.2% of the discrepancies were related to an inconsistency between the practices in community pharmacies and the design of the application (e.g. putting new prescriptions on hold when dispensing an old prescription for the same medication); while 21.3% were related to the rule for attribution of the "current" flag in the SQIM.

Two thirds (63.9%) of patients had inaccurate medication lists generated from the SQIM. To construct appropriate business rules for converting dispensing data from community pharmacies into usable clinical information, a better understanding of community pharmacy work processes is required.

---

[1] Corresponding Author.

# Toward a Virtual Community of Healthcare Facilities: Virtual Lab

Eustach MUTEBA AYUMBA[a][1]

[a]*Independent Researcher, Correspondent of IMIA in DR Congo*

**Keywords.** Virtual Lab, Remote laboratory diagnosis, Malaria.

The proposed solution, a Virtual Laboratory (@Lab), is developed under the project OMAT, Optimization of Malaria Treatment. It is a set of services such as remote microscopy access, remote biomedical image processing, instant-messaging which can be used in a social network dedicated to health professionals interested in issues related to malaria.

The concern of the study is to develop and implement an automatic laboratory diagnosis of malaria. The @Lab (http://www.maesoft1.co/ilab/remotelab.html) is a system that has been designed to acquire, from the microscope, a digital representation of a pathology slide (jpg format) by a digital camera and to view it remotely (as video).

The biomedical image processing of a digital image of the patient's specimen is based on three main features that characterize malaria parasites: colour, form and size. The colour can be used to detect the presence of parasitaemia. The infected cells appear with dark colour. The form and size can be useful to identify malaria type and stage. The identification is done by comparing selected regions of interest of the query image to basic patterns representing common forms of Plasmodium available in the database.

Assessment has been carried out between two workstations of our Research Lab. One workstation was composed by a digital camera with a computer system and the other was composed by a computer system only.

We evaluate the accuracy of the system, with some predetermined level of confidence, to detect the presence of parasitaemia and to identify the Plasmodium species. The identification was limited only to the Plasmodium falciparum species. Also we test the system for collaborative discussion of cases.

The results indicated that the @Lab system is relatively good at the detection of parasitaemia phase. But at the identification phase, the system meets problems to discover patterns and mosaics within a virtual slide caused sometimes by the preparation process of specimen and by the limitation of our artefacts of calculation.

Utilizing our system in developing countries where there is disparity between urban and rural areas can benefit to healthcare practitioners to have a remote access expertise or support.

---

[1] Corresponding Author.

# Statistical Classification of Dyscalculia Symptoms using Smartphones and Behaviour of Magnocellular Cells

Greger SIEM KNUDSEN[a,1]
[a] *Department of Information Science and Media Studies, University of Bergen, Norway*

**Keywords.** Dyscalculia, Classification, Diagnostics Assistance, Application

Having the sufficient expertise to assist pupils diagnosed with dyscalculia, a disability in learning arithmetic, can be challenging for teachers and parents. Therefore, with the ongoing technological development, several forms of computer software are developed for this group. However, the majority of this software is exclusively focused on learning mathematics, not assisting establishing diagnosis. The current project is developing a mobile application for smartphones for assisting diagnosing of dyscalculia, using the magnocellular theory of developmental dyslexia. Users will have constant access to self-evaluation and overview of their own progress.



*Figure 1. Spatial ability evaluation task example*

The application is designed to satisfy user needs by utilizing a set of key functionalities. On the user's side, these are represented as evaluating performance, applied in diagnosis classification using machine learning, and infographics generation of results. On the educator's side, results are received directly from the application through e-mail. The solution is implemented as a high fidelity prototype developed for the Android operating system, targeting the age group of 14-30 years. Figure 1 shows an example of a task evaluation design for spatial ability, one of the parameters for the symptom classifier. Contrast sensitivity, counting, subtilizing and arithmetic are examples of other evaluated skills. The data gathered will be stored and assessed automatically, providing a full overview of the progress.

Expected outcomes includes a greater, more accurate insight into the presence or absence of a correlation between information transmission in magnocellular cells and dyscalculia. For the users, the application is expected to increase access to evaluation and provide more direct feedback, which will not require visit at the educator's office. Educators can also receive evaluation results automatically. An advantage of this approach is improved validity of response data, which is not affected by any confounding factors, such as the educator's or parents' presence. User interface is simple and easy to use to accommodate all age groups.

---

[1] Corresponding Author.

# Internet of Things & Personalized Healthcare

Floor SIEVERINK[a,1], Liseth SIEMONS[a], Annemarie BRAAKMAN-JANSEN[a], and
Lisette VAN GEMERT-PIJNEN[a]

[a] *Centre for eHealth and Wellbeing Research, Department of Psychology, Health &*
*Technology, University of Twente, Enschede, the Netherlands*

**Keywords.** Big Data, Personalized Medicine, Internet of Things, Wearables

The rapid and ongoing digitalization of society leads to an exponential growth of both structured and unstructured data from Internet of Things (IoT) wearable's. This wealth of data opens the door towards the development of automated and personalized real-time feedback systems. However, aren't there any restrictions when using this data as input for data-driven feedback systems? If this is not the case from a technological or semantic point of view, what happens from the ethical perspective? Can we simply collect and connect all the data provided by wearable's and patient health records without questioning, in order to increase the match between the applications, its users and context of use?

Information about IoT usage provides new knowledge about how large and unstructured data sets can be used to improve the usability and persuasiveness of technologies and to personalize the coaching of patients. Current findings in research indicate there is a gap between collecting data and "interpreting and translating" this large data sets into user-friendly, safe, unobtrusive and sense-making feedbacks for patients. To estimate the relevance of the outcomes of data-analysis, a better understanding is needed of models that drive the algorithms to analyze big data.

To better understand the implications of the IoT for our healthcare systems, we conducted a focus group meeting. In this meeting, we investigated the vision, experiences and future ideas of six big data experts from six different scientific disciplines (psychology, philosophy, computer science, business administration, law, and data science) regarding factors that are crucial for using IoT data to support healthcare. We used the 5V model for defining big data as a theoretical framework to categorize the results and to get a multifaceted picture of how to analyze, interpret, and visualize large and complex datasets in an effective, efficient, secure and safe way.
     The experts raised several topics that were arranged in six distinct overall categories: people-driven values, empowerment, profiling, technical infrastructures, and scientific and societal implications. The poster will present a more in-depth overview of the results of the focus group meeting. The findings of these studies will be used to design real-time, accurate, persuasive and personalized feedback systems.

---

[1] Corresponding Author

# Fetching Connected Pedometer Data to Analyze Patients Walking

Julien SOUVIGNET [a,b,1], Claire LE HELLO [c], Béatrice TROMBERT-PAVIOT [a],
Benoît BROUARD [d], Angela CHIEH [d] et Christian BOISSIER [c]

[a] *Service de Santé Publique et de l'Information Médicale, CHU Hôpital Nord, Saint-Etienne, France*
[b] *INSERM, U1142, LIMICS, F-75006, Paris, France; Sorbonne Universités, UPMC Univ. Paris 06, UMR_S 1142, LIMICS, F-75006, Paris, France; Université Paris 13, Sorbonne Paris Cité, LIMICS, (UMR_S 1142), F-93430, Villetaneuse, France*
[c] *Service de Médecine Vasculaire et Thérapeutique – CHU Hôpital Nord, Saint-Etienne, France*
[d] *Withings, Département Santé et Recherche Médicale, Issy-les-Moulineaux, France*

The World Health Organization recommendation is 10,000 steps a day. The generalization of connected pedometers has turned this theoretical number into a practical objective. Pedometers have proven useful to promote a more active lifestyle in the global population. They can also give a concrete answer to the need for remote monitoring of patients. For example, with arteriopathic patients, medical benefits of walking mainly rely on the number of daily episodes of extended walk.

The aim of this study is to obtain accurate data of the spontaneous daily activities of patients using a connected pedometer.

Patients will be equipped with a Withings' "Pulse O2" for a period of 7 days. This device can monitor steps, elevation, distance, calories or heart rate. These figures are linked to a user account and can be fetched through a dedicated Application Programming Interface (API). The API provides multiple services such as "get intraday activity" which allows to recover walking data (up to 1440 records per day). We developed a dedicated Web tool in PHP to fetch API data. To detect continuous episodes of walking, data samples from daily activities are evaluated in term of Metabolic Equivalent of Task (METs) reflecting energy consumption. Each walking episode is then identified by detecting consecutive samples with similar METs value.

Initial tests with 4 users showed that we were able to accurately track trampling periods, periods of slow walking, and sustained walking periods. The Web tool has been validated and can easily fetch and data export.

Connected pedometers store quantity of useful information in term of medical data. Their analysis may reflect daily walking habits, highlight any patient handicap, and offer the potential of a personalized re-training walking program for each patient, whereas, until now, they only had general advice.

---

[1] Corresponding Author.

# Consensus on Norwegian Archetypes

Gro-Hilde ULRIKSEN[a1], Rune PEDERSEN[a,b]

[a] *Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway*
[b] *Telemedicine and eHealth research group, Arctic University of Tromsø, Norway*

In Norway, a new EPR system based on openEHR architecture and archetypes is under development. Since most Norwegian hospitals will use the new EPR it was decided to develop archetypes at a national level by letting them undergo a consensus process to ensure interoperability and cross-organizational EPR usage. Participants were recruited by the National Administration Office of Archetypes (NRUA) to review archetypes by using the web based Clinical Knowledge Manager (CKM) that work as a national repository and a tool for consensus work on archetypes. Since January 2014, national consensus has been reached on 25 archetypes in Norway. There are 126 more in progress and by the end of 2016, 200 archetypes are expected to be ready for use.

Our research includes all data of a single archetype consensus process on the evaluation archetype **problem/diagnosis**. This archetype is used to register persistent data concerning problems and diagnosis identified by the clinician. Clinicians then get access to persistent and structured information. This archetype has been approved both in Norway and internationally through coordinated processes. This is a contribution to a longitudinal qualitative iterative study in the North Norwegian Health region, in coordination with NRUA. The main data sources are the logs of the CKM, in addition to observations, open-ended interviews, and document analysis.

Preliminary findings: The review process to ensure the quality and usability of the translated text both due to semantics and clinical content. To include all necessary participants in the process towards final consensus. The extensive time spent to reach consensus was challenging considering redundancy in addition to fulltime work. There were some important matters brought into the Norwegian consensus process necessary to bring further to the international hearings. For instance, as commented from a Norwegian GP and taken into consideration internationally; *""summarized" and "details" in the same sentence seems illogical. This archetype is supposed to include all details? Not just a summary"*.

The archetype problem/diagnosis went through five iterations in 6 months, 41 participants were involved in the process. Three contributed in all iterations and the average participation rate were two iterations. Important topics discussed in the consensus process were: how to include terminology, language challenges due to translation and semantics, solving clinical issues, and how participants are involved in this process in relation to homogeneity of the group and time use.

---

[1] Corresponding Author.

# Empowering Patients to Choose Appropriate and Safe Hospital Services

Dimitrios ZIKOS[a,1], Karthick SELVARAJ[a], Vivek VAIDYANATHASUBRAMANI[a]
and Pooja PANDEY[a]
*[a] University of Texas at Arlington, USA*

In a consumer driven healthcare system, patients seek access to high quality hospital services. To respond to this demand, we developed a data driven web portal (MediQoC), to provide to Medicare patients (47% of hospital admissions in the U.S) personalized information about their hospital stay, and quality indicators.

The portal utilizes a de-identified, one-million-tuple Medicare in-hospital claims file which contains records for the 100% of Medicare beneficiaries in Texas, United States. The dataset includes admission and discharge information, clinical outcomes, hospital procedures, diagnoses and cost of care.

Users can make a search, either based on their symptom or a clinical procedure they plans to undergo and is presented with a list of National Providers, which have been admitting at high frequency patients with similar profiles. Users can click on any provider to review detailed, personalized information about the hospital stay characteristics, as well as a series of quality indicators, including Hospital Acquired Conditions, in hospital mortality, prevalence of a series of complications, and other. For any selected hospital, the system calculates a unique ranking for the desired quality of care indicator. All parameters can be reviewed on a map module.
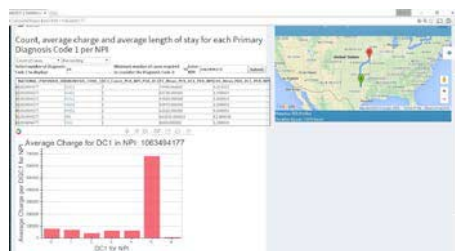


**Figure 1.** The MediQoC interface

Use of medical claims in decision making systems can contribute to better decisions for patients and their caregivers and our strategy is to eventually make MediQoC available to the Medicare patient community. Wide use of platforms as such, which extend and use additional datasets, can eventually leverage the constant efforts for quality and safe hospital services provided by healthcare providers.

---

[1] Dimitrios Zikos, CSE Department, University of Texas at Arlington email: dimitriszikos@gmail.com

# Subject Index

*Transforming Healthcare with the Internet of Things*
J. Hofdijk et al. (Eds.)
135

# Author Index